Noise Estimation, Noise Reduction and Intensity Inhomogeneity Correction in MRI Images of the Brain

Michael Eziashi Osadebey

A Thesis
in
The Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy (Electrical & Computer Engineering) at
Concordia University
Montréal, Québec, Canada

June 2015

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Michael Eziashi Osadebey**

Entitled: **Noise Estimation, Noise Reduction and Intensity Inhomogeneity Correction in MRI Images of the Brain**

and submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (Electrical & Computer Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. E. Shihab _____ Chair

Dr. H. Yahia _____ External Examiner

Dr. W. F. Xie _____ External to Program

Dr. A. Ben Hamza _____ Examiner

Dr. D. Qiu _____ Examiner

Dr. Nizar Bouguila _____ Thesis Supervisor

Approved by _____
Dr. Abdel R. Sebak
Chair of Department or Graduate Program Director

June 15, 2015 _____
Dr. Amir Asif
Dean of Faculty of Engineering & Computer Science

ABSTRACT

Noise Estimation, Noise Reduction and Intensity Inhomogeneity Correction in MRI Images of the Brain

Michael Eziashi Osadebey, Ph.D.
Concordia University, 2015

Rician noise and intensity inhomogeneity are two common types of image degradation that manifest in the acquisition of magnetic resonance imaging (MRI) system images of the brain. Many noise reduction and intensity inhomogeneity correction algorithms are based on strong parametric assumptions. These parametric assumptions are generic and do not account for salient features that are unique to specific classes and different levels of degradation in natural images.

This thesis proposes the 4-neighborhood clique system in a layer-structured Markov random field (MRF) model for noise estimation and noise reduction. When the test image is the only physical system under consideration, it is regarded as a single layer Markov random field (SLMRF) model, and as a double layer MRF model when the test images and classical priors are considered.

A scientific principle states that segmentation trivializes the task of bias field correction. Another principle states that the bias field distorts the intensity but not the spatial attribute of an image. This thesis exploits these two widely acknowledged scientific principles in order to propose a new model for correction of intensity inhomogeneity.

The noise estimation algorithm is invariant to the presence or absence of background features in an image and more accurate in the estimation of noise levels because it is potentially immune to the modeling errors inherent in some current state-of-the-art algorithms. The noise reduction algorithm derived from the SLMRF model does not incorporate a regularization parameter. Furthermore, it preserves edges, and its output is devoid of the blurring and ringing artifacts associated with Gaussian and wavelet based algorithms. The procedure for correction of intensity inhomogeneity does not require the computationally intensive task of estimation of the bias field map. Furthermore, there is no requirement for a digital brain atlas which will incorporate additional image processing tasks such as image registration.

This thesis is dedicated to my jewel of inestimable value, Rebecca Queen Osadebey, to show my appreciation for her enduring patience and emotional support throughout the duration of the PhD study, and to Almighty God, the source of all knowledge and my guide in this earthly journey.

ACKNOWLEDGEMENTS

Special thanks to Prof. Douglas Arnold, the chief executive officer (CEO) of NeuroRx Research Inc., and Marie-Josee Rho, the chief operating officer (COO) of NeuroRx Research Inc., for their moral support throughout the duration of my PhD program.

I would also like to express my special gratitude to the management of NeuroRx Research inc. and the Alzheimer's Disease Neuroimaging Initiative for providing magnetic resonance images of brains as test data for my research.

My appreciation would not be complete without acknowledging the contributions of my course lecturers: Prof. Tal Arbel of McGill University for statistical computer vision, Prof. Omair Ahmed for probability and stochastic processes and Prof. MNS Swamy for engineering analysis. Finally I would like to show my appreciation to my supervisor Prof. Nizar Bouguila and members of the PhD supervisory committee.

NeuroRx Research (`http://www.neurorx.com/en/home.htm`) is dedicated to working with the pharmaceutical industry to facilitate clinical trials of new drugs for multiple sclerosis (MS) and other neurological diseases. NeuroRx provides the professional management of all MRI-related study activities and promptly delivers precise MRI outcome measurements that are performed in a regulatory compliant environment. The organization specializes in logistics of scan handling and tracking and can provide this service exclusively, if needed. NeuroRx uses advanced image analysis techniques to provide precise outcome data that can maximize study power. Images are corrected for inhomogeneity and co-registered for perfect re-alignment and increased precision. Analysis is conducted in 3D, rather than on slices, so that information can be properly related to structures that span multiple slices. Customized automatic segmentation techniques are combined with expert supervision to maximize the precision of outcome measures related to both lesional and non-lesional pathology as well as brain volume changes.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) `(adni.loni.usc.edu)` was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects, but ADNI has been followed by ADNI-GO and ADNI-2. To date, these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see `www.adni-info.org`.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

1.1 Background

1.1.1 Characteristics of MRI Images

The magnetic resonance imaging (MRI) system acquires images in diffferent levels of axial, sagittal and coronal imaging planes. A MRI image at a specific level of an imaging plane is known as slice. A MRI slice potentially has a high spatial and contrast resolution that efficiently discriminates different anatomical structures within a physiological system. These features make MRI system a popular imaging modality for diagnostic procedures and clinical research in the field of medicine. MRI slice image in its original form of acquisition is a complex valued Gaussian distributed data having both real and imaginary components. A combination of the real and imaginary components that produce magnitude MRI nonlinearly transforms Gaussian distribution of pixels in each complex plane into Rician distribution [1].

1.1.2 Characteristics of Brain MRI Images

MRI images of human brains will exhibit unique characteristics. They are generally piecewise smooth and statistically simple [2]. Three structures: white matter, grey matter and cerebro-spinal fluids dominate intra-subject MRI slices. There is inter-subject tissue geometric similarity across age, gender and race [3]. Corresponding inter-subject MRI slices reveal strong spatial, structural and voxel statistical similarities. These similarity features are exploited in the design of MRI-based brain atlases, which are widely acknowledged powerful tools in the analysis of brain images [4], [5], [6].

1.1.3 Brain MRI Images in Clinical Trials

MRI images are an important component of clinical trials of drugs for the treatment of multiple sclerosis, Alzheimer's and other neurodegenerative diseases [7], [8], [9]. Daily, several thousands of brain MRI slice images are delivered from clinical trial sites around the globe to clinical research organizations that manage clinical trials of new drugs for pharmaceutical organizations. Efficient management of this large amount of data demands a high level of automation from image processing and image analysis software at study management centers. The Performance of an image analysis system is strongly dependent on the quality of the MRI image. The acquisition of high quality MRI images are dependent on the optimal performance of MRI device components and optimal selection of MRI system operating parameters. In routine acquisitions these conditions are often not met and will result in degraded MRI images. Noise and intensity inhomogeneity, otherwise known as the bias field are two common types of degradation suffered by MRI images.

1.1.4 Noise

Noise is the random variations in voxel intensity levels where there should be none. Noise changes pixel configurations, weakens and can even obliterate what is supposed to be sharp edges, and can cause low spatial resolution as well as low contrast within and between anatomical structures. These negative features reduce the medical utility of an image because of a difficulty in the study of brain anatomy and the detection of disease signatures. A perfectly acquired image as viewed by the human eye contains some amount of noise. Noise becomes a nuisance when its presence attracts any level of poor visual quality. Noise is unavoidably present in images because of the imperfections of device components and trade-offs in the operating parameters of imaging systems: signal-to-noise ratio (SNR), resolution and length of scan time [10]. A MRI image with very high SNR can be acquired by increasing acquisition time. In practice, acquisition time is compromised for patient comfort, particularly for very young and elderly subjects [11].

1.1.5 Intensity Inhomogeneity

Bias field is the characterization of a specific tissue class or homogeneous anatomical structure by non-uniform voxel intensities instead of uniform voxel intensities. The utility of a MRI image for manual and automated diagnosis can be significantly reduced by the presence of bias field. Visual inspection of a MRI scan degraded by the bias field is more challenging to a physician. It requires more time to discern between the presence of an abnormal anatomical structure and artifact. State-of-the-art image analysis algorithms, which are based on voxel intensities, perform poorly by misclassifying tissue classes because different intensity levels are assigned to the same tissue

class. Several factors are responsible for the bias field. They include the non-uniform sensitivity of radiofrequency coils, non-uniformity of static field, improper acquisition parameter settings, gradient induced eddy currents and magnetization properties of an anatomic structure under investigation [12], [13], [14].

1.1.6 Bias Field Compared to Rician Noise

Bias field and Rician noise have different spatial variations within an MRI image. At different levels of severity, Rician noise is the uniform distribution of a single random variable, but the bias field is the random distribution of at least one uniformly distributed random variable. The distribution and severity of the bias field varies with MRI system magnetic field strength, optimality of selected acquisition parameters, individual patient anatomy and each MRI slice within the volume data. As the magnetic field strength of a MRI system increases, the spatial variation of a bias field changes from locally smooth inhomogeneity to globally non-smooth inhomogeneity [15], [16], [17].

1.2 Problem Statement

The quality of a MRI image for analysis reflects the performance of noise reduction and bias field correction units within an image processing system. It is therefore necessary to carry out noise reduction and intensity inhomogeneity correction before manual or automatic image analysis. The noise reduction unit preserves sharp edges that define boundaries of anatomical structures and disease signatures. A bias field correction unit assigns similar intensity levels to the same tissue class.

Noise estimate is required as an input parameter for algorithms developed for the task of noise removal [18], [19], [20], segmentation [2], [21], registration [22], quality assessment of functional magnetic resonance images [23], overall quality assessment of magnetic resonance imaging systems [24], [25] and the performance evaluation of noise removal algorithms [26].

In the field of medical imaging and computer vision, noise reduction is at a crossroad between image processing and image analysis. Its invasive nature makes it the most delicate fundamental problem because there is a high risk of introducing extraneous features that can further degrade image quality and reverse gains derived from previous image processing tasks. Noise reduction algorithms are expected not only to reverse the effect of noise but also to optimize the performance of image analysis systems by preserving edges and other features that determine the level of clinical utility of the image.

In the past thirty years researchers made several contributions on the correction of the bias

field in medical images. Despite the significant contributions, correction of intensity inhomogeneity is still an active research area because of increasing clinical interest on brain MRI images and the development of high field MRI systems that produce more complex characteristics of the bias field [27], [28]. None of the current algorithms address the problem of correcting the different characteristics of intensity inhomogeneity across different levels of magnetic field strength [29].

Many noise estimation, noise reduction and intensity inhomogeneity correction algorithms are based on strong parametric assumptions. They adopt generic models that do not account for class-specific salient features in natural images and different characteristics of a particular image degradation. Thus, it is not unexpected that there is yet no algorithms that exhibit optimal performance for different characteristics of a particular image degradation, different imaging modalities and for all classes of natural images. Their limitations manifest as high computational cost and difficulty in the estimation of model parameters that encourage heuristics and manual tuning in estimation of model parameters. The limitations also include an inaccurate estimate of noise levels, introduction of extraneous artifarts such as blurring, staircasing and spurious edges into restored images

1.3 Contributions

This thesis focuses on optimizing the quality of a particular class of natural images such as the magnetic resonance images of the brain, for efficient image analysis. It exploit simplicity and similarity of statistics of human anatomy for different subjects in a random experiment to propose a two-in-one algorithm that adopts Markov random field energy as metric for the estimation of noise variance in brain MRI. The first approach casts MRF energy as a variable in a random experiment. The second formulates the MRF energy-noise level relationship as a mathematical model. Both approaches are fast, accurate and efficient in the estimation of noise variance. It is invariant to presence or absence of background features in an image and is potentially immune to modeling errors inherent in current state-of-the-art algorithms.

The concept of Variational-Bayesian (VB) cycle was introduced to propose a new computationally efficient approach for the selection of a regularization parameter in a TV denoising algorithm. The thesis also propose a new general framework, built from the 4-neighborhood clique system, for denoising MRI images of the brain. The kernel quantifies the smoothness energy of spatially continuous anatomical structures. Scalar and vector valued quantification of smoothness energy configures images for Bayesian and variational denoising modes, respectively. Within variational mode, the choice of norm adapts images for either total variation or Tikhonov technique. The proposal integrates Bayesian and variational techniques into a two layer-structured denoising

system for optimal performance in specific applications. The first layer of the new denoising system is a hitherto unknown form of Markov random field model. This new model is referred to as single layer Markov random field (SLMRF) model because it denoises images by minimizing energy but without reference to a classical prior image model. The new algorithm is simple and computationally efficient as it does not incorporate a regularization parameter.

Another proposal is a new and novel approach to correct intensity inhomogeneity in brain MRI. An anatomic structural map regarded as a pseudo digital brain atlas, generated directly from the test image, guides the algorithm for automatic operation. Accurate information from the structural map is combined with the distorted intensity level attributes of the test image to detect outliers in regions of interest (ROI) generated by k-means clustering. The number of ROI is cleverly chosen to equal the number of tissue classes in k-means clustering. Outliers in each ROI are merged with voxels in the appropriate tissue class. Intensity levels of the new set of voxels in each tissue class are rescaled to conform to intensity levels of uncorrupted voxels. Performance evaluation demonstrates the efficiency of this proposal for different characteristics of the bias field.

1.4 Outline of Thesis

This thesis is organized as follows: Previous contributions on noise estimation, noise reduction and intensity inhomogeneity correction are reviewed in chapter 2. Chapter 3 describes the theory behind this research, followed by the SLMRF model for noise estimation in chapter 4. Chapter 5 is on the SLMRF model for image denoising. A general framework for image denoising is presented in chapter 6. Intensity inhomogeneity correction is presented in chapter 7. Chapter 8 contains the conclusion.

Chapter 2

Literature Review

2.1 Noise Estimation

Eight contributions on the estimation of noise levels in brain MRI are considered and classified in chronological order as earlier and later contributions.

2.1.1 Earlier Contributions

The earliest, referred to as double acquisition method [30], acquires the same image twice, align and subtract. Noise level is estimated from the standard deviation of subtracted image. Other early contributions [31], [32], [33] estimate noise from the background region where noise is described by Rayleigh distribution.

2.1.2 Later Contributions

The first of later contributions [34] estimates noise level automatically from the maximum likelihood estimate of the partial histogram of a noise signal in the background region. Second in same category [35] estimates noise from local statistics in the foreground where voxel intensities contaminated by noise are described by Gaussian distribution at a high signal-to-noise-ratio and Rician at low signal-to-noise ratio. The third [36] adopts the maximum likelihood estimation principle to estimate the noise level from local variance and local skewness of voxels. The last contribution [37] transforms the image into the frequency domain and estimates noise using median absolute deviation (MAD) in sub-band where image structures are considered suppressed.

2.1.3 Limitations of Current Noise Estimation Algorithms

Double Acquisition

The required correct alignment of images in the earliest contribution is not easily realized. Also, the technique cannot be adopted in a clinical diagnosis requiring a single MRI acquisition experiment such as when contrast material-enhancement is injected into a patient and in functional studies [38].

Wavelets

Wavelet methods overestimate noise at low levels because they revolve around the original thresholding method postulated in [39], [40]. The authors in [37] attempt to address this problem through an application of correction factor proposed in [41] to noise described by Rician distribution.

Earlier and Later Contributions

Modern MRI machines are designed with chambers filled with perfluorinated fluids to stabilize unwanted movements of patients and to eliminate artifacts. Their images have a signal-free background. Expectedly, techniques in both classes [31], [32], [33], [34], [35] that exploit the presence of background information will be obsolete as they will be useless in applications where background signals are absent. They are prone to significant error where background information is limited or corrupted by artifacts [36].

2.2 Noise Reduction

The contributions from each author on image denoising presented as a unique technique can be contained in several volumes of literature articles. Careful review reveals that sophisticated and state-of-the-art techniques can be classified into two major classes: global and specific schemes. Global schemes are a general framework for specific schemes, and their formulation generally results in a new specific scheme. A geometrical framework proposed by Sochen et al. [42] combines linear heat flow, anisotropic diffusion and mean curvature flow techniques into a single scheme. The deformation by curvature technique proposed by Kimla and Siddiqi [43] unifies geometric heat equation and anisotropic diffusion. Hamza et al. [44] propose the unification of probabilistic and variational techniques, and Auclair-Fortier and Ziou [45] propose the global approach for solving the heat equation. The detailed review of image denoising methods can be found in [46]. For this review only six techniques were considered.

2.2.1 Gaussian Filters

Earliest and classical technique is the use of linear and non-linear filters [47], [48], [49]. They are designed based on the Gaussian assumption of noise distribution and cast on a moving window over the noisy image in a convolution process. In the classical Gaussian filtering, a single two-dimensional Gaussian kernel K_σ:

$$K_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp - \left(\frac{|x|^2+|y|^2}{2\sigma^2}\right) \tag{2.1}$$

of standard deviation σ is cast on a window moving in two orthogonal directions u, v, over the noisy image $I_d(x,y)$ of size MXN, in a convolution process:

$$K_\sigma(x,y) * I_d(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} I_d(x-u,y-v)K_\sigma(u,v) \tag{2.2}$$

2.2.2 Anisotropic Diffusion

The convolution process stated in Eq. 2.2 is the solution of the two dimensional diffusion equation of a physical system [50], [51], [52]:

$$\frac{\partial}{\partial t} I_d(x,y,t) = \nabla.\left(D(x,y,t)\nabla I_d(x,y,t)\right) \tag{2.3}$$

where the physical system under consideration is an image $I_d(x,y,t)$ over time t with diffusion tensor D and initial condition $I_d(x,y,0) = I_d(x,y)$. In the operation of classical Gaussian filter, the diffusion tensor is a constant value, $D = 1$, independent of x, y or t, and the physical system is said to exhibit homogeneous diffusivity, resulting in equal smoothing and hence blurring in all directions. To reduce blurring, the diffusion tensor is configured as a symmetric positive-definite diffusion tensor:

$$D = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{2.4}$$

The eigenvectors V_1, V_2 and eigenvalues λ_1, λ_2 are chosen in such a way as to encourage smoothing in homogeneous regions and discourage smoothing in the direction of edges. This is the basis of the anisotropic diffusion method pioneered by Perona and Malik [53].

2.2.3 Non-Local Means

An alternative to the Perona and Malik method is to determine a denoised voxel $I(i)$ at position i in a neighborhood Ω_i of an image from the Gaussian weighted linear combination of all neighborhood

Ω_j of voxels j outside Ω_i:

$$I(i) = \frac{1}{C} \sum_{j \in \Omega} w(i,j) I(j) \tag{2.5}$$

where C and $w(i,j)$ are normalization constants and similarity weights, respectively.

$$C = \sum_{j \in \Omega} w(i,j) \tag{2.6}$$

The similarity features are proportional to measures by a Gaussian kernel of standard deviation a in a square neighborhood G:

$$w(i,j) \propto \exp\left(-\frac{\|G_i - G_j\|^2_{2,a}}{h^2} \right) \tag{2.7}$$

The degree of filtering by the Gaussian kernel is controlled by h. This is the idea behind the non-local means algorithm introduced by Buades and his collaborators [46], [54].

2.2.4 Wavelets

Wavelet approach [39], [10] is a multiresolution analysis. Wavelet function is the family of wavelets generated by scalings a and translations b of a mother wavelet $\Psi(t)$ [55]:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \Psi\left(\frac{t-b}{a} \right), \quad a,b \in \mathbb{R}, \quad a \neq 0 \tag{2.8}$$

In an image processing task such as denoising, a family of wavelet is convolved with a degraded image I_d to generate wavelet coefficients $c_{a,b}$. The image is represented as a linear combination of wavelets coefficients

$$I(t) = \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} c_{a,b} \Psi_{a,b}(t) \tag{2.9}$$

As in Fourier transform, the image is reconstructed from wavelet coefficients. In the transformed domain image signals are sparse, wavelet coefficients are assumed to be Gaussian distributed, and noise signal is spread out equally along all coefficients. These properties are exploited to remove noise by reconstructing the image from coefficients by retaining and eliminating the coefficients of image that are stronger and insignificant respectively relative to energy γ of the Gaussian noise.

This technique is known as wavelet soft T^S and hard T^H thresholding method:

$$T^S(c_{a,b}) = \begin{cases} 0 & \textbf{if } |c_{a,b}| \leq \gamma \\ c_{a,b} - \gamma & \textbf{if } c_{a,b} > \gamma \\ c_{a,b} + \gamma & \textbf{if } c_{a,b} < -\gamma \end{cases} \tag{2.10}$$

$$T^H(c_{a,b}) = \begin{cases} 0 & \textbf{if } |c_{a,b}| \leq \gamma \\ c_{a,b} & \textbf{if } |c_{a,b}| > \gamma \end{cases} \tag{2.11}$$

The method was pioneered by Donoho and Johnstone [39]. The threshold is a function of noise level σ and number of voxels N in the image:

$$\gamma = \sigma\sqrt{2\log N} \tag{2.12}$$

Noise level is determined by the median absolute deviation of wavelet coefficients at the finest resolution level. Other thresholding methods have been formulated by manipulating parameters in Eq. 2.10, Eq. 2.11 and Eq. 2.12. They include VISU shrink [56], SURE shrink [57], [58] and Bayes shrink [59].

2.2.5 Markov Random Fields

In 1984, Donald Geman and Stuart Geman introduced Markov random field (MRF) theory [60] to the computer vision community. They proposed the MRF model, within the Bayesian framework, for image restoration. The underlying image is assumed to exhibit spatial coherence and is hence modeled as a Markov random field with a known probability distribution. The observed image, referred to as likelihood, given any realization of an underlying image, is assumed to follow a Gaussian probability distribution. These probabilities are cast into the Bayesian framework followed by the optimization of the posterior probability distribution [61]. A denoised image is an image that maximizes the posterior probability referred to as MAP. Alternatively, the denoising process can be cast as a constrained optimization of the energy functional.

2.2.6 Total Variation

Total variation technique [62], [63], exploits the knowledge that noise introduces spurious details into clean images resulting in a high absolute gradient. The denoising process is a constrained minimization in which the absolute gradient is minimized subject to a similarity with an underlying prior image model.

2.2.7 Limitations of Current Noise Reduction Algorithms

Gaussian Filters

At low noise levels, when image structures dominate over noise signals, Gaussian filters in a single operation are excellent noise filters on images with voxels that are Gaussian distributed. At a significant level of noise, when image structures compete for visibility alongside noise signals and multiple operations of denoising process are required, its rotationally invariant property results in smoothing both edges and homogeneous regions [64]. This reduces the utility of the image in medical diagnosis. It is from this characteristic that the filter earns the name of smoothing kernel [65].

Total Variation

In the image processing community, total variation (TV) is widely acknowledged as a popular and state-of-the-art technique for noise reduction because of its edge-preserving property. This attractive feature of TV is dependent on the optimal selection of a regularization parameter. Contributions in literature on TV focus on applications, properties and different numerical solution methods. The few contributions [66], [67] that address the problem of regularization parameter selection are based on regression methods which pre-exist the introduction of TV. This parameter is easily achieved by manual tuning for a single image. In MRI-based clinical trials management centers, where the daily routine includes the processing and analysis of several thousands of brain MRI images, current TV regularization parameter selection techniques will be computationally inefficient [68].

In specific applications, the need to understand the application environment for optimal performance calls for the manual tuning of a parameter until the best peak signal-to-noise ratio (PSNR) of the image is attained [69], [70], [71]. The design and operation of different regularization parameter algorithms incorporate heuristics. For example in generalized cross validation (GCV), minimum number of observations required for optimal performance has to be heuristically determined from the experiment by the user. Noise will prevail in the regularized solution if a reasonable number of data points are not chosen as input into the algorithm. Moreover, GCV does not have a universal application as there are reports of failure in some model parameter selection problems [72]. Performance of the discrepancy principle, the unbiased predictive risk estimator (UPRE) and Stein's unbiased risk estimator (SURE) relies on the good estimate of an error level, otherwise there is risk of over-smoothing or retaining noise [73].

Anisotropic Diffussion

Experimental results have shown that overcoming blurring in the anisotropic diffusion technique is an ideal expectation because the regularization approach to achieving well-posedness of the diffusion model introduces blurring into the denoised image [74], [50]. Though a suitable diffusion coefficient may be chosen to serve the purpose of edge preservation for a particular image, the choice may not be optimal when a large number of images are processed. Also, the iterative nature of the process often introduces blurring in flat regions of the image.

Wavelets

Finding an optimal threshold in wavelet technique is a challenging task as reflected by volumes of articles devoted to thresholding methods. Local structures of the image and wavelet coefficients at each scale cannot be well described by Gaussian distribution. The result is blurring and ringing artifacts [75], [76].

Non-Local Means

The non-local means technique is widely considered as a modern, effective and popular denoising technique because it does not introduce artifacts as in wavelets, but it is known to be computationally inefficient and reduces the sharpness of edges as is found with filters. Computational inefficiency is the result of calculating similarity weights in the full-space of neighborhood [77]. Reduction of sharp edges is the effect of calculating similarity weights in the presence of noise [78]. Its performance is dependent on the accurate estimate of noise level in the image. Interestingly, several contributions in literature that highlight these limitations made proposals to address the problems [79], [80], [77], [78].

Markov Random Field

At inception, MRF technique was regarded as a revolutionary technique in image denoising. A serious setback is the difficulty in estimating model parameters of the probability distribution. This encourages a significant manual effort in the denoising process [81]. Another setback is the computational inefficiency of algorithms for the computation of MAP, such as iterative conditional modes [82], simulated annealing [83] and loopy belief propagation [84]. Though graph cuts [85] was introduced as a much faster and efficient algorithm, the level of improvement does not make it a popular choice for denoising large volumes of data commonly found in clinical trials management centers.

2.3 Intensity Inhomogeneity Correction

2.3.1 Design Techniques of Current Algorithms

Current bias field correction schemes are based on a common mathematical model. They can be classified into two classes, prospective and retrospective, according to the object of interest for the source of information in the elimination of the bias field.

2.3.2 Mathematical Models of the Bias Field

The mathematical model that explains the bias field in MRI system is derived from the generally accepted mathematical model describing imaging process in different imaging modalities [86], [13]. At location index l within an image consisting of L number of voxels, a voxel with observed intensity value I_l and the underlying uncorrupted signal x_l are related according to:

$$I_l = b_l x_l + \varepsilon \tag{2.13}$$

where b_l and ε are a spatially varying function of the image coordinate denoting the bias field and uniformly distributed random noise due to the imaging device. To account for signal variation within a specific tissue class j the general model in Eq. 2.13 is reformulated as:

$$I_l = b_l(u_j + \varepsilon_j) + \varepsilon \tag{2.14}$$

The expression in Eq. 2.14 shows that the signal emitted by an anatomical structure belonging to a specific tissue class is a random variable corrupted with noise ε_j and distributed around a mean value u_j. The biological noise ε_j arising from a specific anatomical structure and noise ε due to the MRI device makes the determination of the bias field in Eq. 2.14 nontrivial. The model yields equations with intractable solutions when different anatomical structures are under consideration [87]. To simplify the model, the following two assumptions are made on the bias field and the imaged object [88], [89].

(a) The bias field is a slowly varying variable that can be approximated by a constant within a clearly defined neighborhood.

(b) The image is piecewise constant and can be approximated as the sum of all disjoint regions.

For computational convenience, the intra-tissue variability represented by the term $(u_j + \varepsilon_j)$

is neglected so that the general model is rephrased [90], [86]:

$$I_l = b_l u_j + \varepsilon \tag{2.15}$$

Another mode to computational convenience is to account for the intra-tissue signal variation and assume that the bias field is independent of image noise, so Eq. 2.14 becomes:

$$I_l = b_l(u_j + \varepsilon_j) \tag{2.16}$$

The computational complexity of Eq. 2.16 is further reduced through a logarithm transformation, which turns the multiplicative bias field b into an additive bias field [91], [92]:

$$\log I_l = \log b_l + \log u_j + \log e_j \tag{2.17}$$

Correction of intensity inhomogeneity, preceded by noise removal, is a two-step process. First is the estimation of intensity inhomogeneity in the image domain followed by division in Eq. 2.13, Eq. 2.14, Eq. 2.15 and Eq. 2.16 or subtraction in the log domain in Eq. 2.17.

Prospective Method

In prospective methods the imaging device is the object of interest because it is regarded as the source of intensity inhomogeneity. The bias field is corrected by adopting correct acquisition parameters and acquisition protocols. This can be achieved by calibrating the MRI system with phantoms, special imaging sequences and the combination of two different radio frequency coils [93], [94].

Images of objects with uniform physical properties such as oil and water are used as phantoms. Intensity inhomogeneity of the image is estimated from the filtered and smoothed MRI image of the phantom [94]. The combination of surface and volume coils allows each individual coil to compensate for the deficiency of the other coil. A surface coil characterized by good signal-to-noise-ratio (SNR) but vulnerable to intensity inhomogeneity is compensated for with presence of a volume coil which has poor SNR but is resistant to intensity inhomogeneity [95].

Retrospective Method

Retrospective methods have the acquired image as an object of interest and correct the bias field using *a prior* knowledge and properties measured from an acquired image. There are several categories of retrospective method. In this thesis we review only four popular categories. They are low frequency filtering, surface fitting, segmentation and histogram:

1. *Low Frequency Filtering*: The low-pass filtering technique exploits the slowly varying property of the bias field. The bias field and image details are assumed to be localized in the lower and upper range of frequency domain, respectively. This property enables the extraction of the bias field from a corrupt image. Homomorphic filtering, popular for nonuniform illumination correction, is in this category. This filtering method follows the mathematical model in Eq. 2.17 by computing the bias corrected image as the difference between the log transformation of a degraded image and the log transformation of a low frequency filtered image [96]. A similar algorithm, homomorphic unmask filtering corrects the bias field according to the mathematical model in Eq. 2.16 [94].

   Wavelet decomposition is another low pass filtering approach to correct intensity inhomogeneity. The degraded image is decomposed into approximate space and residual space. An approximate template in approximate space is regarded as consisting of multi-scale sensitivity profiles of a surface coil of the MRI system. The sensitivity profile that gives optimal output from a suitably chosen filter reflects the sensitivity profile of the surface coil. Bias field correction is achieved using the optimal sensitivity profile [97], [98].

2. *Surface Fitting*: A surface fitting method fits parametric smooth functions such as polynomials and splines to image features that characterize slowly varying intensity inhomogeneity. Image features may be intensities [99] or local gradients of intensities [100], [101]. A popular surface fitting method is Parametric Bias field Correction referred to as PABIC. It is based on a parametric model of tissue statistics and a polynomial model of intensity inhomogeneity field. The estimation of the bias field is formulated as non-linear energy minimization [90].

3. *Segmentation*: Segmentation method is based on the principle that accurate segmentation trivializes intensity inhomogeneity correction. It is widely acknowledged that alternating segmentation and intensity inhomogeneity steps will optimize output of an image analysis task. Examples of segmentation methods include fuzzy c-means (FCM) [91], mixture models [102], [103] and level set [88].

4. *Histogram*: Histogram methods manipulate the intensity histogram of the image for the correction of the bias field. There are three popular techniques under histogram based bias field correction methods.

   The nonparametric nonuniformity normalization (N3) is in this category. Design of N3 is based on a multiplicative mathematical model and assumes that the histogram of an image represents the probability distribution of a given signal. It corrects the bias field by seeking a bias field that maximizes the frequency content of the histogram of an image [104].

Another approach in this category considers that intensity inhomogeneity increases the information content of an image. Entropy, a measure of information content, is extracted from the image intensity histogram and cast in the information minimization framework to correct the bias field [105].

Yet another approach seeks to manipulate histogram information in a local neighborhood of image where intensity inhomogeneity is assumed to be constant. Local neighborhoods of image are obtained by dividing the image into subvolumes. The intensity inhomogeneity map is estimated by comparing the histogram in each locality of the test image to a corresponding locality in a histogram model [106].

2.3.3 Importance of Current Algorithms

The several volumes of research articles on the correction of the bias field reflect the importance of brain MRI in the study and examination of neurodegenerative diseases. The old and popular homomorphic filtering approach is very simple to implement and can operate in real-time. The design of FCM makes it robust to the partial volume effect. Iteration between segmentation and bias field correction in segmentation based techniques improves the accuracy of the corrected bias field. Histogram based methods are suitable for automated operations because prior knowledge of image and initialization is not required as in segmentation based methods. The detailed specific contributions of different algorithms for correction of intensity inhomogeneity is in [13].

2.3.4 Limitations of Current Algorithms

Limitations of current algorithms are inherent in their design formulation or derived from adopted mathematical models. These limitations are categorized as algorithm and model limitations.

Algorithm Limitations

Phantom based methods cannot correct for patient induced inhomogeneities. Both phantom based methods and combination of surface and volume coils for bias correction require more acquisition time, which is often impracticable in large scale studies such as clinical trials [16], [13]. Frequency spectrums of homogeneous regions that describe specific anatomical structures are in the same low range as bias fields. There is a high risk of eroding relevant anatomical structure when the filtering method of intensity homogeneity correction is applied [107]. The blind extrapolation of intensity inhomogeneity from a major tissue to the entire image in an intensity based surface fitting method limits the accuracy of bias field correction. Gradient based methods in the same class as a surface

fitting method can give an erroneous bias field estimate if the homogeneous image regions are not large enough or are dominated by abnormal tissue such as the presence of a tumor [13]. There is a risk of introducing extraneous artifacts or eroding anatomical structures because many current bias field correction techniques are not robust to images corrupted by little or no bias field [29].

Model Limitations

Since 1971, when the MRI system was invented, increased clinical interest in brain MRI encouraged the development of high magnetic field strength MRI systems. Operating magnetic field strength increased from 1.5T to 7T for humans and as much as 9.4T and 14.1T for animals. The use of a higher magnetic field improves the signal-to-noise ratio and acquisition time, but the bias field, if present, can be significant, severe and complex [27]. The common assumption of a smoothly varying multiplicative bias field adopted by most current algorithms is only an ideal expectation to make the problem of bias field correction tractable. Experience has shown that, even with this assumption, the bias fields computed by most of these algorithms such as [14] and [104] are in general not smooth. Extra resources and sometimes heuristics approach is adopted to maintain smoothness [15]. In the past, some researchers had cast doubt on the reliability of adopting a smooth multiplicative intensity inhomogeneity field as a model of the bias field [108]. The doubt was made strong when the contribution of [109] demonstrates that intensity inhomogeneity field is not necessarily globally smooth. The field may be locally smooth but not globally smooth. In such situations most of the current algorithms will fail to correct the bias field.

Chapter 3

Theoretical Background

The theory behind this doctoral research and thesis is anchored on three concepts: the 4-neighborhood clique kernel, layer-structured Markov random field (MRF) model and set theory. The 4-neighborhood clique system, which has its origin in graphical models [110], defines the relationship between the intensity of a pixel and that of its neighbors within an image. Layer-structured MRF model is rooted in the Markov random field model [111]. In layer-structured MRF model, images under consideration are regarded as layers. When the test image is the only physical system under consideration, the MRF model is regarded as single layer. It is regarded as double layer when the test image and a classical prior image are considered.

This chapter begins with the description of a 4-neighborhood kernel. Next is the description of a single layer Markov random field (SLMRF) model and how the assignment of different smoothness costs to a 4-neighborhood kernel determines scalar or vector valued total clique potential of an MRI slice image. The Rudin, Osher and Fatemi (ROF) total variation technique [62], Tikhonov regularization technique [112] and the classical Markov random field model [111] are classified as double layer Markov random field models. At end of the chapter there is explanation on how to apply set theory [113] as an important step in the algorithm for the correction of intensity inhomogeneity.

3.1 The 4-Neighborhood Clique System

Given a 2D image I of size $m \times m$ and $M = m \times m$, a four neighborhood clique system $\mathbf{N}$ having a clique size of two is the set consisting of four pixels that are closest in distance to any pixel located at (x,y):

$$\mathbf{N} = \{\mathbf{N_1} = (x, y+h), \mathbf{N_2} = (x+h, y), \mathbf{N_3} = (x, y-h), \mathbf{N_4} = (x-h, y)\} \tag{3.1}$$

where x and y are indices of rows and columns within the image, **N1**, **N2**, **N3** and **N4** are labels identifying each neighbor, and h is an interval between each grid. The structure of the system is displayed in Fig. 3.1a. The red colored pixel is the reference pixel and the four neighbors are the black colored pixels. The intensity levels of pixels that constitute the image belongs to a family of random variables $F = \{F_1, \ldots, F_M\}$ defined on the set of pixel locations S called sites [111]:

$$S = \{(x,y) | 1 \leq x,y \leq m\} \tag{3.2}$$

The assignment of intensity levels, as in a grey level image, from set $L = \{0,1,\ldots,255\}$ to each site S is an event called configuration $f = \{f_1, \ldots, f_M\}$. In the clique system the reference pixel located at (x,y) is denoted i and any of its neighbors $N1$, $N2$, $N3$ and $N4$ as expressed in Eq. 3.1 is denoted $i^{'}$ and assigned configurations f_i and $f_{i'}$, respectively [111].

3.2 4-Neighborhood Kernel and Gibbs Energy

Configurations assigned to a reference pixel and its neighbors are adopted as indices to compute Gibbs energy. Gibbs energy $U(f)$ of an image with pixel configuration $f \in F$ is the sum of single site V_1 and pair-site V_2 cliques [111]:

$$U(f) = \sum_{i \in S} \alpha_1 V_1(f_i) + \sum_{i \in S} \sum_{i' \in N} \alpha_2 V_2(f_i, f_{i'}) \equiv E_d + E_s \tag{3.3}$$

where the first and second terms are data E_d and spatially varying smoothness E_s terms, respectively, and α_1 and α_2 are interaction coefficients. At each local clique N, where reference pixel is $i \in S$, contribution of each neighboring pixel $i^{'} \in N$ to local smoothness energy is determined as:

$$V_2(f_i, f_{i'}) = \begin{cases} \xi_r & \textbf{if } f_i = f_{i'} \\ \xi_p & \textbf{otherwise} \end{cases} \tag{3.4}$$

where ξ_r is reward and ξ_p penalty for conformity $f_i = f_{i'}$ and violations $f_i \neq f_{i'}$ of smoothness constraints, respectively. Henceforth, the three terms: Gibbs energy, Markov random field energy and smoothness energy are considered equivalent, have the same meaning for a given image and will be used interchangeably. The matrix formulation of a kernel for computing local clique potential energy is:

$$\mathbf{H} = \begin{bmatrix} 0 & \xi_{r,p} & 0 \\ \xi_{r,p} & 0 & \xi_{r,p} \\ 0 & \xi_{r,p} & 0 \end{bmatrix} \tag{3.5}$$

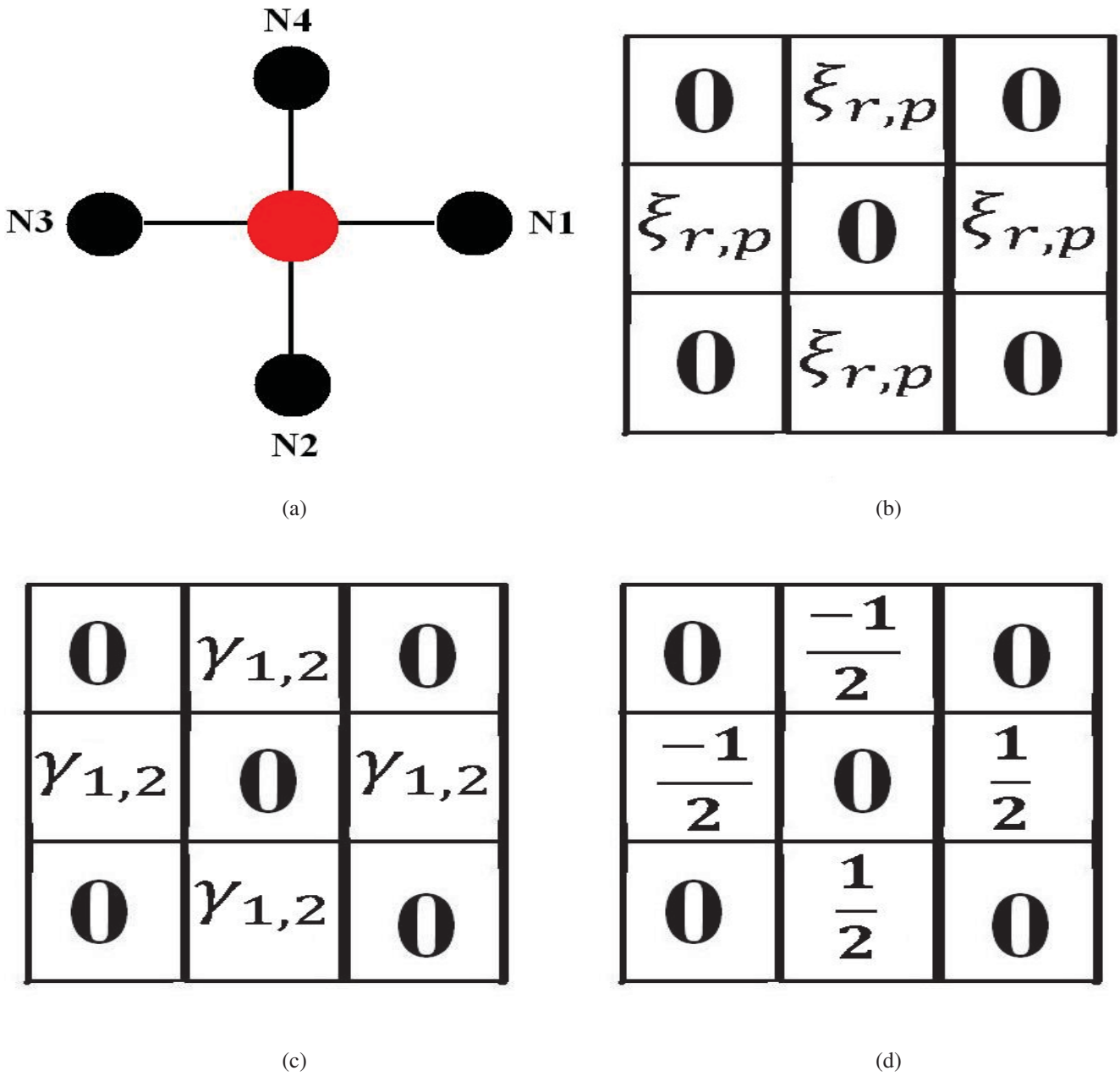


Figure 3.1: (a) Structural description of the 4-neighborhood clique system. The four nearest neighbor of a pixel(in red color) located at (x,y) are the black colored pixels labeled $N1$, $N2$, $N3$ and $N4$. (b) The general form of the kernel for the 4-neighborhood clique system. (c) Kernel for configuring an image for the Bayesian mode of operation. (d) Kernel for configuring an image for the variational mode of operation.

where $\xi_{r,p}$ denotes the assignment of either reward ξ_r if $f_i = f_{i'}$ or penalty ξ_p if $f_i \neq f_{i'}$. The physical structure of the kernel corresponding to this matrix is displayed in Fig. 3.1b.

The assignment of scalar values γ_1 and γ_2 to ξ_r and ξ_p, respectively, in Eq. 3.4 results in a

scalar valued smoothness cost at each local clique:

$$V_2(f_i, f_{i'}) = \begin{cases} \xi_r = \gamma_1 & \textbf{if } f_i = f_{i'} \\ \xi_p = \gamma_2 & \textbf{otherwise} \end{cases} \tag{3.6}$$

and **H** in Eq. 3.5 becomes:

$$\mathbf{H_s} = \begin{bmatrix} 0 & \gamma_{1,2} & 0 \\ \gamma_{1,2} & 0 & \gamma_{1,2} \\ 0 & \gamma_{1,2} & 0 \end{bmatrix} \tag{3.7}$$

where $\gamma_{1,2}$ denotes the assignment of either reward γ_1 if $f_i = f_{i'}$ or penalty γ_2 if $f_i \neq f_{i'}$. In Eq. 3.4, the assignment of vector values of ξ_{r_j} and ξ_{r_k} to ξ_r, and ξ_{p_j} and ξ_{p_k} to ξ_p in orthogonal directions, **j** and **k** results in a vector valued smoothness cost expressed as:

$$V_2(f_i, f_{i'}) = \begin{cases} \xi_{r_j} = \xi_{r_k} = 0 & \textbf{if } f_i = f_{i'} \\ \xi_{p_j} & \textbf{otherwise, j direction} \\ \xi_{p_k} & \textbf{otherwise, k direction} \end{cases} \tag{3.8}$$

Penalties ξ_{p_j} and ξ_{p_k} can be expressed using standard image location notation:

$$\begin{aligned} \xi_{p_j} &= \frac{I(j+h,k) - I(j-h,k)}{2h} \\ \xi_{p_k} &= \frac{I(j,k+h) - I(j,k-h)}{2h} \end{aligned} \tag{3.9}$$

In this case **H** in Eq. 3.5 becomes:

$$\mathbf{H_v} = \begin{bmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \tag{3.10}$$

The physical structure of the kernels corresponding to $\mathbf{H_s}$ and $\mathbf{H_v}$ are displayed in Fig. 3.1c and Fig. 3.1d, respectively.

3.3 Single Layer Markov Random Field

In the single-layer Markov random field, unlike classical Markov random field theory, the observed image is the only physical system under consideration, and there is no reference to a prior model.

Pixel locations are quantized into 4-neigbourhood clique system that describes spatial coherence or clusters of similar pixels. The single layer Markov random field energy is the sum of local clique potentials V_c:

$$U(f) = \sum_{c \in C} V_c(f) \tag{3.11}$$

Gibbs energy U is dependent on pixel configuration f and is computed from the sum of potential function contributions from a single site and pair-site cliques expressed in Eq. 3.3.

Noise changes the pattern of arrangement of pixels in the image as demonstrated in Fig. 3.2. There is a gradual change in pixel configurations for noise levels increasing from $\sigma = 0$ in Fig. 3.2a to $\sigma = 30$, $\sigma = 60$ and $\sigma = 90$ in Fig. 3.2b, Fig. 3.2c and Fig. 3.2d, respectively. Different noise levels result in a change in strength of clusters within the image. Thus, different noise levels σ give rise to different pixel configurations and their corresponding levels of energy $\hat{U}_\sigma(f)$. The data term is set to zero because it is a constant for different levels of noise. Thus, the smoothness energy can be regarded as the total clique potential per pixel U:

$$U(f) = \frac{\hat{U}}{D} = \sum_{i \in S} \sum_{i' \in N} V_2(f_i, f_{i'}) \tag{3.12}$$

normalized by D, the product of row and column dimensions of image.

In the degradation process, different image pixel configurations $f \in F$ can be realized from an observed image $I_{\sigma_n}(f)$ degraded by noise σ_n of level n. The convolution of $I_{\sigma_n}(f)$ with $\mathbf{H_s}$ is the scalar valued total smoothness energy $U(f)$ computed from the sum of local clique potentials. We denote the scalar value as γ and express it as a function of image pixel configuration $\gamma(f)$:

$$U(f) = \mathbf{H_s} * I_{\sigma_n}(f) = \gamma(f) \tag{3.13}$$

On the other hand, convolution with $\mathbf{H_v}$ results in total smoothness energy $U(f)$ that is vector valued and is expressed as a function of image gradient $\nabla I(f)$:

$$U(f) = \mathbf{H_v} * I_{\sigma_n}(f) = \nabla I(f) \tag{3.14}$$

The magnitude of vector valued total energy can be measured either in L_q norm:

$$L_q = (\|\nabla I\|_q) \tag{3.15}$$

where $q = 1, 2$ is choice of norm.

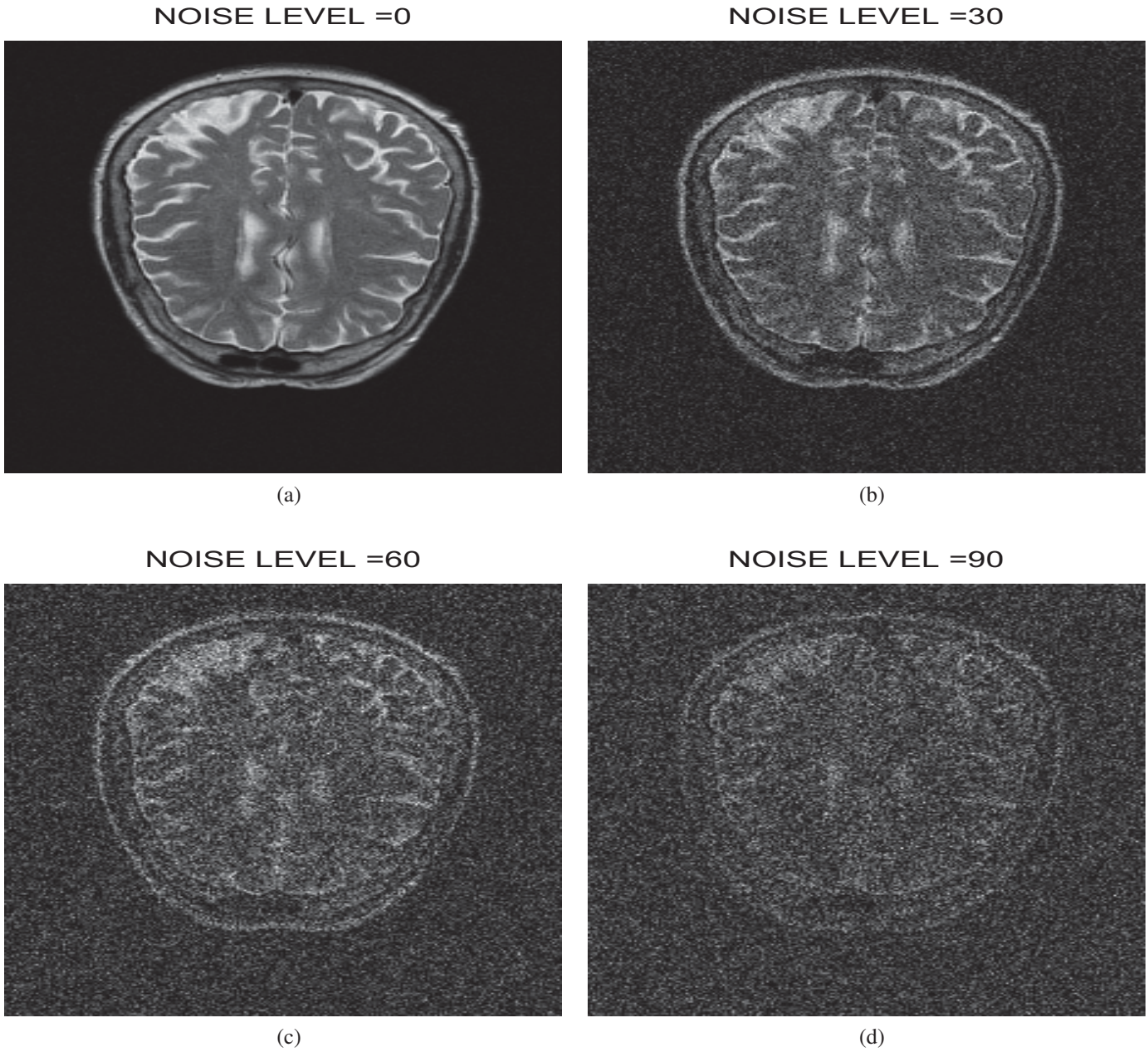


Figure 3.2: (a) An MRI slice image at various levels of degradation by Rician noise. Observe the different image pixel configurations associated with noise levels of $\sigma = 0$, $\sigma = 30$, $\sigma = 60$ and $\sigma = 90$.

3.4 Double Layer Markov Randon Field: ROF Total Variation and Tikhonov Techniques

The Rudin, Osher and Fatemi total variation technique [62] follows the same theoretical formulation as the single layer Markov random field explained in the previous section, but there is another

physical system under consideration. It is an underlying clean image modeled as piecewise smooth and referred to as prior model I_c. The estimate of the true image $\hat{I}_c$ is derived by the minimization of a functional subject to the observed image I_d being closely matched to the prior:

$$\hat{I}_c = \min_{I_{\sigma_n}(f)} \quad \sum_x \sum_y \|(\nabla I_{\sigma_n}(f))_{x,y}\| + \frac{1}{2\lambda}\|I_{\sigma_n}(f) - I_c\|^2 \tag{3.16}$$

where λ is the regularization parameter. If the L_1 norm of vector valued smoothness constraint in Eq. 3.16 is changed to the L_2 norm, Rudin, Osher and Fatemi total variation technique changes to the Tikhonov or Sobolev regularization technique [112]:

$$\hat{I}_c = \min_{I_{\sigma_n}(f)} \quad \sum_x \sum_y \|(\nabla I_{\sigma_n}(f))_{x,y}\|^2 + \frac{1}{2\lambda}\|I_{\sigma_n}(f) - I_c\|^2 \tag{3.17}$$

Parameter $[\lambda : 0 \leq \lambda < \infty\}$ is a positive constant referred to as the regularization parameter. This parameter is a weight that balances the level of piecewise smooth property of a prior image that is impacted to the denoised image and level of noise removal. The parameter is also a measure of fidelity $(I_{\sigma_n}(f) - I_c)$ to the prior image I_c. The selection of λ in three different ranges of values results in denoised images with three different properties:

$$\lambda \quad \Rightarrow \begin{cases} \textbf{Too High} & \hat{\mathbf{I}}_\mathbf{c} \Rightarrow \textbf{denoised \& oversmooth} \\ \textbf{Too Low} & \hat{\mathbf{I}}_\mathbf{c} \Rightarrow \textbf{noisy \& piecewise smooth} \\ = 0 & \hat{\mathbf{I}}_\mathbf{c} = \mathbf{I_d} \quad \textbf{denoised image same as observation} \end{cases} \tag{3.18}$$

If the value of λ tends to be too high, the algorithm is constrained to place more emphasis on noise removal, but the denoised image is oversmooth because there is less emphasis on the impacting piecewise smooth property of the prior image on a denoised image. On the other hand, the value of λ that tends to be too low results in a denoised image that possesses the piecewise smooth property of a prior image but retains noise. A special case is when $\lambda = 0$ in which a denoised image is the same as a noisy image. Thus, the optimal performance of both ROF and Tikhonov techniques is strongly dependent on the optimal selection of λ [73].

3.5 Classical Double Layer Markov Random Field Model

The seminal report in [60] can be divided into four successive steps. In the first step, the pixel configuration f of underlying image I_c is assumed to exhibit spatial coherence and is hence modeled as a Markov random field. Markov-Gibbs energy equivalence is invoked so that the probability

of occurrence of a particular configuration $P(f)$ of an underlying image obeys the Gibbs distribution [111]:

$$P(f) = P(I_c(f)) \propto \exp -(U(f)) \tag{3.19}$$

where $U(f)$, energy function, is the sum of clique potentials $V_c(f)$ over all possible cliques C as expressed in Eq. 3.11. The probability distribution of each realization of the observed image I_d from prior I_c is assumed to follow a Gaussian distribution:

$$P(I_d(f)|I_c) \propto \exp -\left(\frac{(I_d(f)-I_c)^2}{2\sigma^2}\right) \tag{3.20}$$

In the second step each possible image pixel configuration of prior and likelihood are assumed to be independent and identically distributed. Thus, probabilities $P(I_c)$ and $P(I_d|I_c)$ of prior and likelihood are expressed as:

$$P(I_c) \propto \prod_{f\in F} \exp -(U(f)) \tag{3.21}$$

$$P(I_d|I_c) \propto \prod_{f\in F} \exp -\left(\frac{(I_d(f)-I_c)^2}{2\sigma^2}\right) \tag{3.22}$$

In the third step probabilities in Eq. 3.22 and Eq. 3.23 are inserted into Bayes posterior probability $P(I_c|I_d)$ formula [114], [111]:

$$P(I_c|I_d) \propto P(I_d|I_c)P(I_c) \tag{3.23}$$

The fourth step is the optimization of the posterior probability distribution:

$$\begin{aligned} \hat{I_c} &= \arg\max_{I_d} \quad P(I_c|I_d) = \arg\max_{I_d} \quad [P(I_d|I_c)P(I_c)] \\ &= \arg\max_{I_d} \quad \left\{ \left[\prod_{f\in F} \exp -\left(\frac{(I_d(f)-I_c)^2}{2\sigma^2}\right)\right] \left[\prod_{f\in F} \exp -(U(f))\right] \right\} \end{aligned} \tag{3.24}$$

This formulation estimates the denoised image $\hat{I_c}$ from the image configuration that maximizes $P(I_c|I_d)$, referred to as maximum *a posteriori* probability (MAP).

3.6 Set Theory Description of MRI Slice Image Quality

The MRI image is a signal Y defined in an image domain Ω. The signal represents M types of anatomical structures and has three attributes, namely intensity I, spatial location R and entropy S.

Entropy is derived from R. These attributes are in Ω, $\Omega_I \in \Omega$, $\Omega_S \in \Omega$ and $\Omega_R \in \Omega$. There are M disjoint regions in the domain of each attribute:

$$\Omega_I = \{\Omega_{I,1}, ., \Omega_{I,M}\} \tag{3.25}$$

$$\Omega_R = \{\Omega_{R,1}, ., \Omega_{R,M}\} \tag{3.26}$$

$$\Omega_S = \{\Omega_{S,1}, ., \Omega_{S,M}\} \tag{3.27}$$

For each attribute, the regions representing each of the M anatomical structures can be segmented by a suitable image processing technique. Intensity levels attributes is segmented by, for example, the k-means clustering algorithm [115]. Entropy attribute is segmented by combining entropy filtering with thresholding. Spatial location attributes can be segmented by computing the region property across the image.

3.7 An Ideal MRI Image

The following holds true for an ideal MRI image acquired using correct acquisition parameters and acquisition protocol:

$$\Omega = \Omega_I \cup \Omega_R \cup \Omega_S \tag{3.28}$$

$$\emptyset = \cap_{j=1}^{M} \Omega_{I,j} \tag{3.29}$$

$$\emptyset = \cap_{j=1}^{M} \Omega_{R,j} \tag{3.30}$$

$$\emptyset = \cap_{j=1}^{M} \Omega_{S,j} \tag{3.31}$$

$$\emptyset = \Omega_{I,j} \cap \Omega_{R,k}, \quad j \neq k \tag{3.32}$$

$$\emptyset = \Omega_{I,j} \cap \Omega_{S,k}, \quad j \neq k \tag{3.33}$$

3.8 A MRI Image Corrupted by the Bias Field

We make the following eight statements regarding a MRI image corrupted by a bias field:

(1) The MRI signal has been processed and is without noise.

(2) The bias field signal, unlike the noise signal, is not uniformly distributed throughout the image.

(3) Like residual noise signals, there does exist residual bias field signals. They are insignificant and can be neglected since they do not adversely affect image quality and image analysis.

(4) Henceforth, intensity levels not significantly affected by bias fields will be termed clean signals; otherwise they are termed corrupt signals.

(5) The domain defined by the intensity attributes of each tissue class is the union of the set of clean voxels $\{I_j\}$ and the set of corrupt voxels $\{I_k\}$:

$$\Omega_{I,j} = \{I_j\} \cup \{I_k\}, \quad j \neq k, \quad k = 1,,M \tag{3.34}$$

The expression in Eq. 3.34 implies that voxels associated with a bias field in a particular tissue class belong to a different anatomical structure.

(6) The corrupt voxels are in the image locations termed outliers G. Outliers are voxel locations that are members of the set that violates expressions in Eq. 3.29, Eq. 3.30, Eq. 3.31, Eq. 3.32 and Eq. 3.33:

$$\Omega_{I,j} \cap \Omega_{R,k} = G_{R,k}, \quad j \neq k \tag{3.35}$$

$$\Omega_{I,j} \cap \Omega_{S,k} = G_{S,k}, \quad j \neq k \tag{3.36}$$

(7) Both corrupt and clean signals in each tissue class are described by a normal distribution.

(8) The intensity level attribute of an image have a tissue-specific numeric meaning [116] that potentially trivializes the manual and automated analysis of a MRI scan. The presence of a bias field distorts intensity levels and makes vague their meaning with respect to a tissue class. Nevertheless, the spatial attributes of anatomical structures under investigation are not distorted.

Chapter 4

Single Layer Markov Random Field Model for Noise Estimation

This chapter describes a new algorithm for the estimation of noise levels in the magnetic resonance images of the human brain [117]. Estimation of noise level is modeled as a random experiment. Objects of the random experiment are original 3 mm T2-weighted axial brain MRI images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The images had undergone post-acquisition processing and are assumed to be noise-free. We chose as many as 422 subjects to satisfy the requirements of the central limit theorem [118], [119], [120] and to optimally extract inter-subject similarities in voxel statistics. The range, mean and standard deviation of the subjects' ages are (71-87), 75 and 10, respectively.

In the experiment, for each MRI slice, variations in total clique potential (TCP) with noise level is computed. The TCP are indexed by their corresponding level of Rician noise and stored in a database referred to as the TCP-noise level database. Each outcome in the sample space of the experiment is the total clique potentials and corresponding noise level for each MRI slice in the database. The random variable is the index of the total clique potential in each outcome of sample space that is closest in value to the total clique potential measured from a test MRI image. The total clique potential of a test MRI image is measured, normalized and matched with each element of sample space. An estimate of noise level is derived from the mean and standard deviation of probability distribution generated from the matching process. A schematic diagram of the algorithm is displayed in Fig. 4.1.

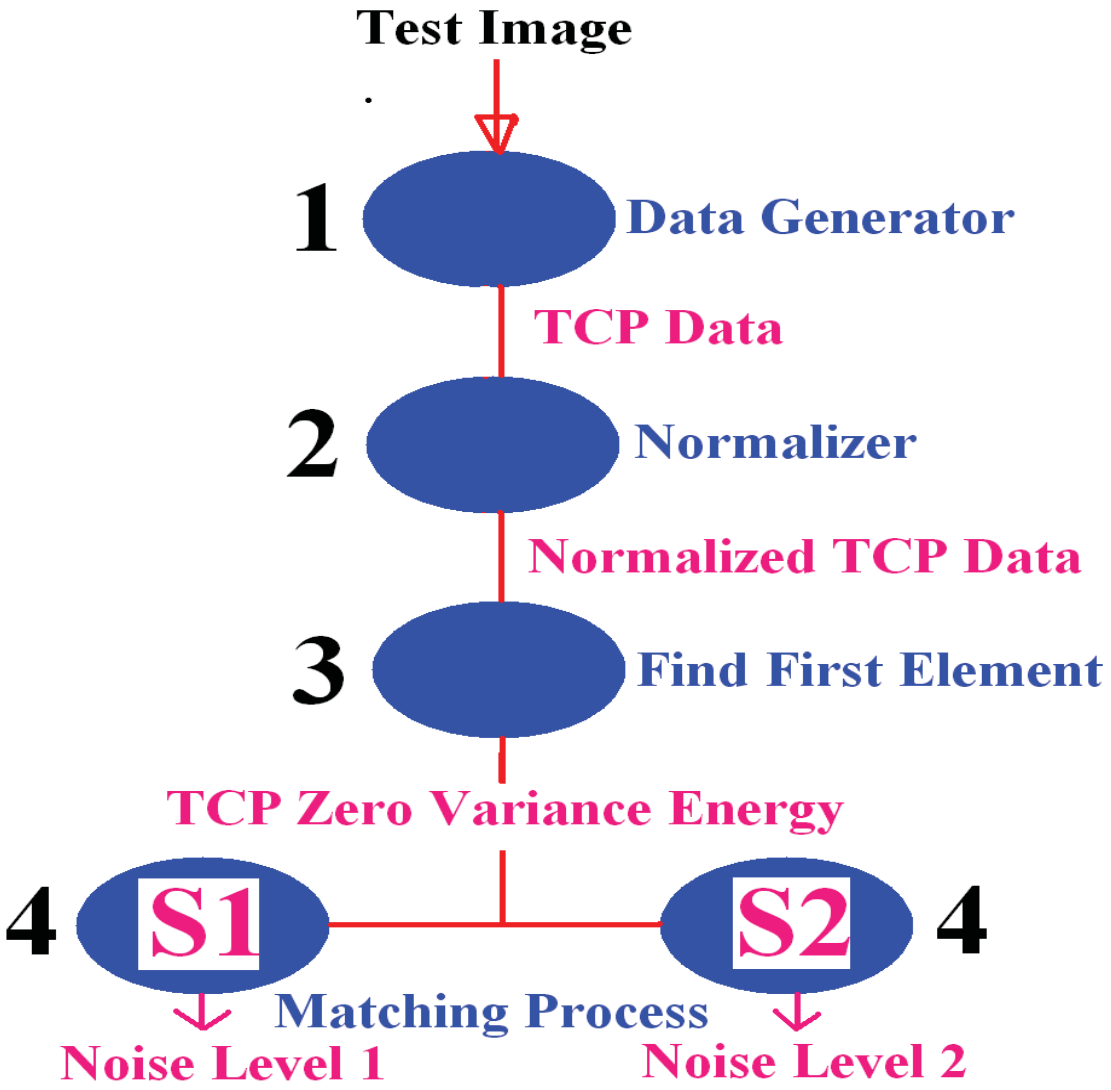


Figure 4.1: The four steps in the estimation of noise level in 2D brain MRI. In step 1, TCP-noise level data of the test image is generated. The data is normalized in step 2 and the first element (TCP zero variance energy) of the normalized data is extracted in step 3. In the estimation of the noise (step 4), the TCP zero variance energy is separately matched to each of the 1D and 2D sample spaces, S1 and S2, respectively.

4.1 Sample Space

Let $W = 422$ be the number of noise-free 3D brain MRI data from subjects across age, gender and race. From the MRI data of each subject, $G = 21$ useful slices are extracted to obtain a total of $Z = GW$ slice images where each MRI slice image is indexed by $z \in [1, Z]$. Each MRI slice image indexed by z is replicated L number of times, and each replica is further indexed by a replica number $l \in [1, L]$. Thus, each MRI slice image $I_{z,l}$ is indexed by its MRI slice number z in the database and its replica number l as defined by the set $\{(z,l) : 1 \leq z \leq Z, 1 \leq l \leq L\}$.

Starting from zero noise level, $\sigma = 0$, each copy of a MRI slice image is corrupted by the increasing level of Rician noise $\{\sigma_l : \sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_L\}$ weighted by the replica index. For each noise level of a replica image, the total clique potential energy (with and without background pixels) $U_{z,l}$ is computed from the sum of local clique potentials according to Eq. 3.3 to form a $1 \times L$ TCP energy-noise level data $\{U_{z,l} : U_{z,1}, U_{z,2}, U_{z,3}, \ldots, U_{z,L}\}$. Each element of the TCP energy-noise level data is normalized by dividing it with the maximum value of the total clique potential in the TCP energy-noise level data so that $(-\infty \leq U_{z,l} \leq 1)$. For the entire MRI slices in the database, we obtain a two dimensional sample space Ω of the normalized variations of TCP energy $U_{z,l}$ with noise level σ_l having dimensions $Z \times L$:

$$\Omega(\omega) = \begin{pmatrix} U_{1,1}, U_{1,2}, \ldots, U_{1,L} \\ U_{2,1}, U_{2,2}, \ldots, U_{2,L} \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ U_{Z,1}, U_{Z,2}, \ldots, U_{Z,L} \end{pmatrix} \tag{4.1}$$

where each outcome ω of the sample space is each row of the matrix defined by $\omega_z = \{U_{z,l} : U_{z,1}, U_{z,2}, U_{z,3}, \ldots, U_{z,L}\}$. For notational convenience we adopt the same notations for normalized and unnormalized TCP data.

A second sample space κ having one dimension of size $1 \times L$ is derived from the two dimensional sample space by computing the mean of Eq. 4.1 along its column:

$$\kappa(l) = \frac{1}{Z} \sum_{z=1}^{Z} U_{z,l} \quad \forall \quad l \tag{4.2}$$

Using the regression analysis we describe κ by a power model:

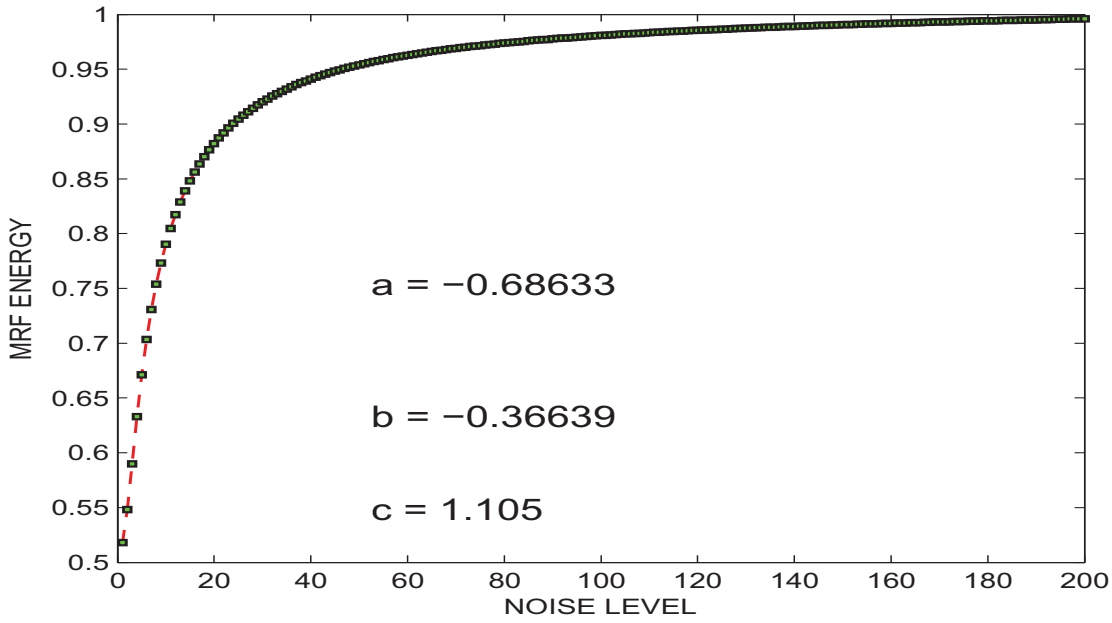
$$E = a\hat{\sigma}^b + c \tag{4.3}$$

$$a_b = -1.67, a_f = -0.6863 \tag{4.4}$$

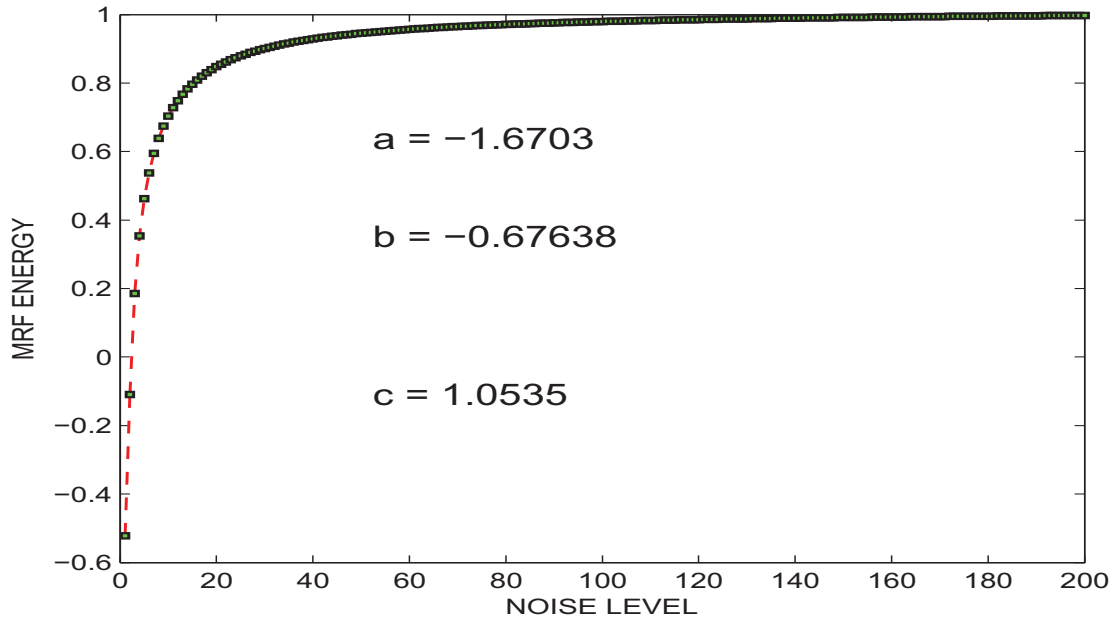
$$b_b = -0.6764, b_f = -0.3663 \tag{4.5}$$

$$c_b = 1.053, c_f = 1.105 \tag{4.6}$$

where $a_b, a_f, b_b, b_f, c_b, c_f$ are model parameters for foreground (f) and background (b) modes, respectively. Plots of the power models are shown in Fig. 4.2a for foreground and Fig. 4.2b for background modes.



(a)



(b)

Figure 4.2: The plots of the proposed generalized mathematical models for describing the relationship between TCP energy and noise level in the (a) foreground and (b) background modes.

4.2 Random Variable

A random variable X is defined as a function that assigns a real number $X(\omega)$, $\omega \in \Omega$, to each outcome ω in the sample space Ω of a random experiment [118]. Different random variables

can be defined based on this random experiment. For example, there are a total of $T = ZL$ MRF energies indexed by $\{L : L < Z\}$ levels of noise. If Z is large enough, such that $Z \gg L$, then T MRF energies will be randomly distributed into each noise level. Thus, the set of all possible MRF energies $\{\widehat{U}_{1,l}, \widehat{U}_{2,l}, \widehat{U}_{3,l}, \cdots, \widehat{U}_{Z,l}\}$, consisting of Z number of elements that are indexed by a particular noise level l is a random variable. The random variable of interest to this proposal is defined as the set containing the indices $X_z(\omega) = \{l_1, l_2, l_3, \cdots, l_Z\}$, $l \in [1, L]$ of TCP energy $U_{z,l}$ in each outcome $\omega_z = \{U_{z,l} : 1 \leq z \leq Z, 1 \leq l \leq L\}$ of the sample space that is closest in value to the TCP energy measured from a test image:

$$\{X_z(\omega)\}_{X_z \in [1,L]} = \arg \min_{l \in [1,L]} : \{l \mid \forall z : \|E_t - \{U_{z,l}\}\|\} \tag{4.7}$$

4.3 Estimation of Noise Level

The test image for evaluation was provided by NeuroRx Research Inc. It was corrupted with an increasing level $\{l : 0 \leq l \leq 200\}$ of Rician noise, starting from the first noise level, $\sigma = 0$. For each noise level, the TCP energy $E_{t.l}$ is computed according to Eq. 3.3 to obtain a 1×200 TCP energy-noise level data. The data is normalized by dividing each element with the maximum TCP energy. The normalized TCP energy $\widehat{E}_{t,1}$ measured from the test image is the first element $l = 1$ of the normalized TCP energy-noise level data.

The normalized TCP energy measured from the test data is matched separately to each TCP energy-noise level data of an MRI slice in the 1D and 2D sample spaces. In both cases we find indices of TCP energy in each TCP energy-noise level data that are closest in value to the normalized TCP energy measured from the test data.

First matching generates a random variable consisting of a sequence of indices $X_z(\omega) = \{l_1, l_2, l_3, \cdots, l_Z\}$ of the same number of elements Z as MRI slice images in the sample space giving a total of Z number of Q unique discrete indices $\theta_z \in [0, L]$ each with frequency of occurrence λ_q. Estimated level of noise $ENV1$ is the mean of probability distribution:

$$ENV1 = \theta_1(\frac{\lambda_1}{Z}) + \theta_2(\frac{\lambda_2}{Z}) + \theta_3(\frac{\lambda_3}{Z}) \ldots + \theta_Z(\frac{\lambda_Q}{D}) - 1 \tag{4.8}$$

Second matching generates only a single and unique index number $ENV2$ because matching is with a single MRI energy-noise level data:

$$ENV2 = \min_{l \in [1,L]} \|\widehat{E}_{t,1} - \kappa_l\| - 1 \tag{4.9}$$

Subtraction of unity in both Eq. 4.8 and Eq. 4.9 is the rescaling of indices to account for the index

of noise level associated with Rician noise of zero variance.

4.4 Comparative Performance Evaluation Results

The two methods of estimating noise in the proposed algorithm are denoted **ENV1** and **ENV2**. They were compared, based on mean square error and absolute mean square error, alongside five current algorithms, namely [31] which we denote **BRU**, [32] denoted **CHA**, [34] denoted **SJI**, [35] denoted **AJA** and [36] denoted **RAJ**. Each slice in the real MRI data was modified to follow Rician noise distribution ranging from $(0 \leq \sigma \leq 40)$. For each original noise level σ the estimated level of noise $\widehat{\sigma}$ from an algorithm is determined from the mean value of the noise level estimated for all the individual slices in the volume data.

4.4.1 Mean Square Error

The second proposal for estimation of noise level **ENV2** demonstrates superior performance in terms of mean absolute error (≈ 0.08) on T2 (Fig. 4.3a) and FLAIR (Fig. 4.3d) MRI images. In the same evaluation parameter, **AJA** was the best algorithm with (≈ 0.07) for T1 (Fig. 4.3b) and PD (Fig. 4.3c) images. It is closely followed by **RAJ** and the proposed **ENV2**.

4.4.2 Root Mean Square Error

In the plot for the root mean square error displayed in Fig. 4.4, the proposed **ENV2** stands out clearly as best for T1 (Fig. 4.4b), (≈ 0.02) and T2 (Fig. 4.4a), (≈ 0.2) images. Closely behind is the proposed **ENV1**, **AJA** and **RAJ**. The algorithm **AJA** is best for PD images (Fig. 4.4c), (≈ 0.2) closely followed by the two proposed noise estimation algorithm, **ENV1** and **ENV2** and **RAJ**. As can be seen in (Fig. 4.4d) for FLAIR images, the first approach in the proposal **ENV1** and **AJA** show the same level of best performance (≈ 0.2) and is closely followed by the second proposal and **RAJ**.

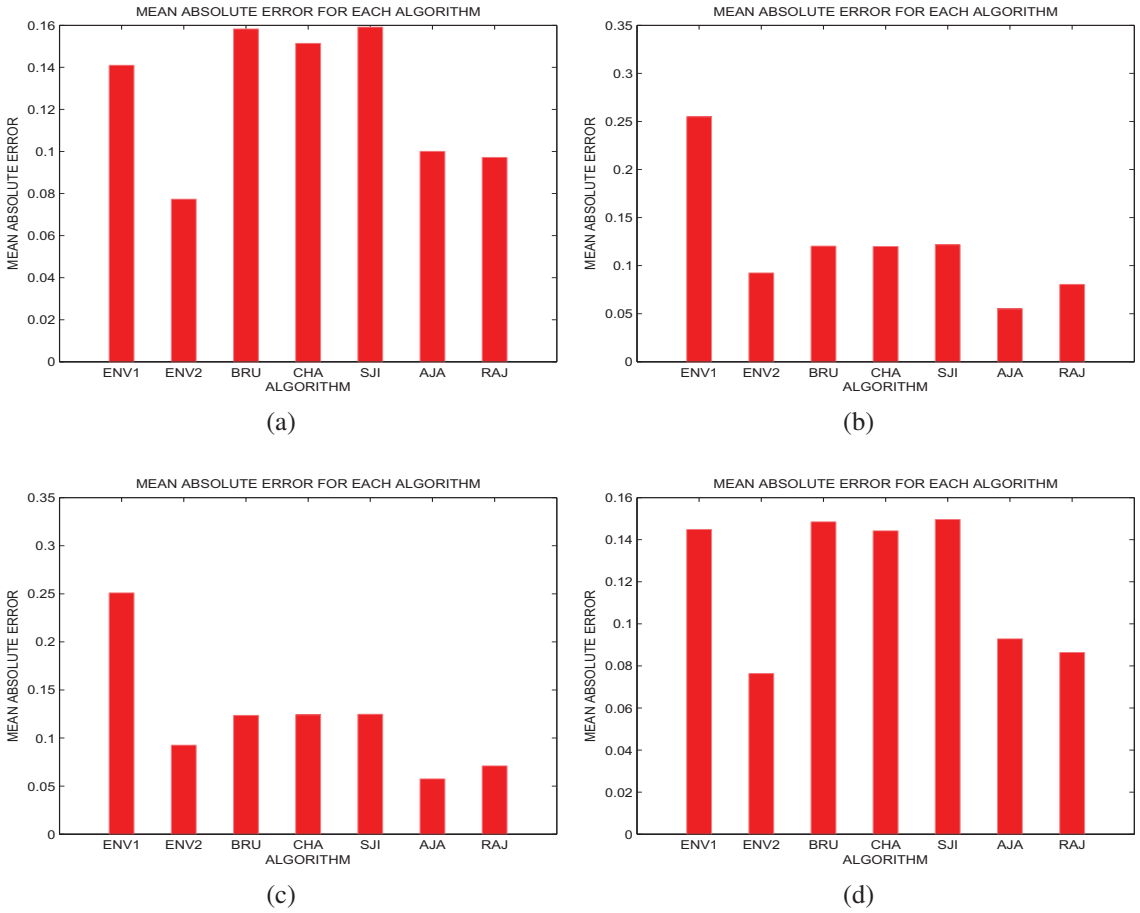


Figure 4.3: Comparative performance evaluation. The plots of the mean absolute error for four different MRI acquisitions - (a) T2, (b) T1, (c) PD and (d) FLAIR.

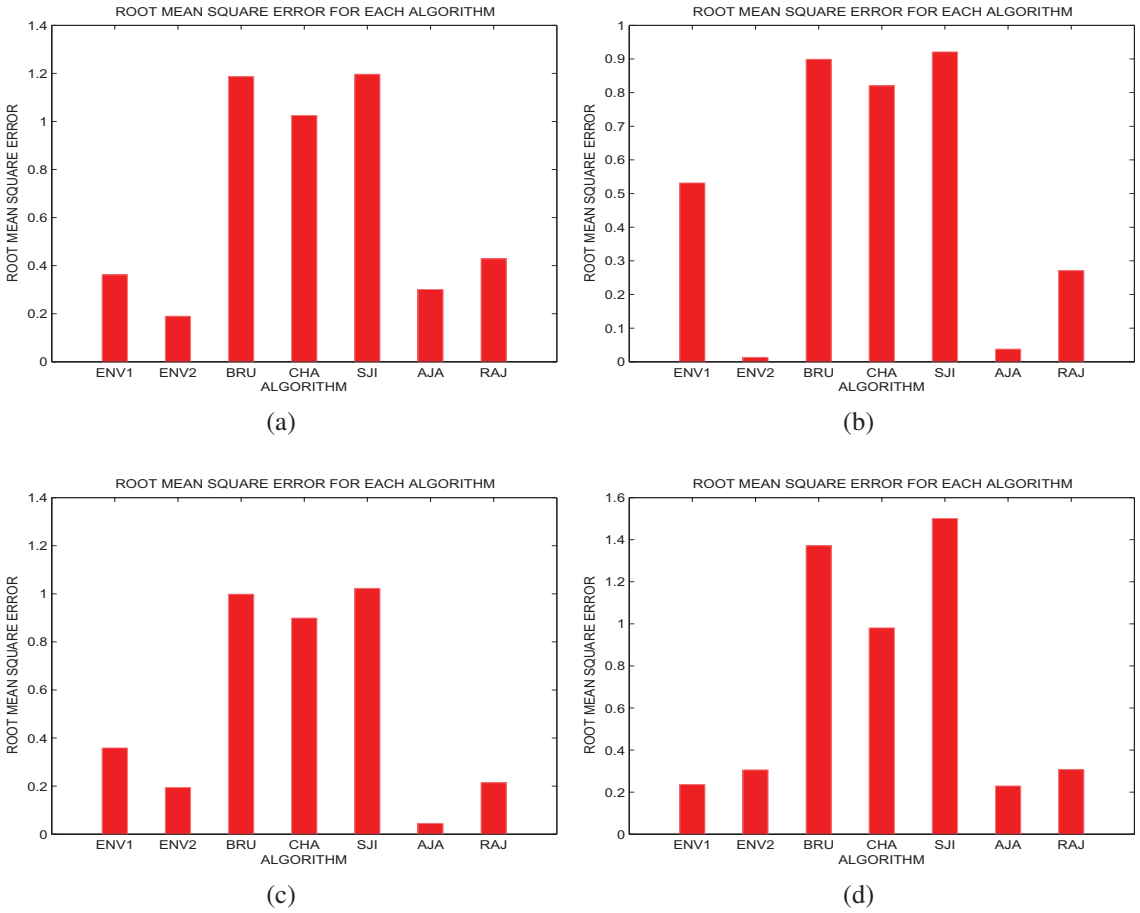


Figure 4.4: Comparative performance evaluation. The plots of the root mean square error for four different MRI acquisitions - (a) T2, (b) T1, (c) PD and (d) FLAIR.

Chapter 5

Single Layer Markov Random Field Model for Denoising

This chapter presents a method of reducing noise using the single layer Markov random field model [121]. In the single layer Markov random field model for reducing noise, the observed image belongs to a specific class of natural images such as a brain MRI slice image $I_{\sigma_n}(f)$ having pixel configuration f degraded by noise σ_n of level n. The observed image is the only physical system under consideration, and there is no reference to a classical prior image model. The chapter begins with explanation on the relationship between total variation, smoothness energy and noise level of an image and how these variables are cast in Euler-Lagrange (E-L) differential equation framework. The solution to the E-L equation is stated followed by determination of the denoised image using gradient minimization scheme. A comparative performance evaluation is at the end of the chapter.

5.1 Total Variation, Smoothness Energy and Noise Level

The image is assumed to be initially clean and without noise $\sigma_0 = 0$ as shown in Fig. 3.2a. The energy at the noise-free state is regarded as the equilibrium energy state. The observed noisy or higher energy state σ_n is attained in a step-wise incremental $\Delta\sigma$ degradation by noise. Different levels of noise $\sigma_{30} = 30$, $\sigma_{60} = 60$ and $\sigma_{90} = 90$ results in different pattern of arrangement of pixels, different spatial coherence and hence different image energy as shown in Fig. 3.2b, Fig. 3.2c and Fig. 3.2d. Different levels of noise σ_n result in different pixel configurations $I_{\sigma_n}(f)$, different image gradients $\nabla I(\sigma_n(f))$ and hence different levels of energy $U_n(\sigma_n(f))$. Thus, the smoothness energy U of image I is a functional $U(\sigma_n, I_{\sigma_n}(f), \nabla I(\sigma(f)))$ having three variables namely image variance σ_n, image pixel configurations $I_{\sigma_n}(f)$ and image gradient $\nabla I(\sigma(f))$ measured in L_1 norm.

Noting that smoothness energy is equivalent of total variation or the integral of absolute gradient $(\sum_x \sum_y (\|(\nabla I_{\sigma_n}(f))_{x,y}\|))$ of the image, the smoothness energy can be expressed as:

$$U(\sigma_n, I_{\sigma_n}(f), \nabla I(\sigma(f))) = \sum_x \sum_y (\|(\nabla I_{\sigma_n}(f))_{x,y}\|) \tag{5.1}$$

The task of noise reduction is to return the image from higher energy state σ_n to the equilibrium energy state σ_0.

5.2 Euler-Lagrange Differential Equation

The Euler-Lagrange partial differential equation [122], [123] corresponding to Eq. 5.1 is :

$$\frac{\partial U}{\partial I} - \frac{\partial U}{\partial \sigma} \frac{\partial U}{\partial (\|\nabla I\|)} = 0 \tag{5.2}$$

The classical method to derive the solution of Euler-Lagrange equation is by adopting the finite difference method of approximating partial derivatives [124]. To make the expression in Eq. 5.2 well-posed the variable $\|\nabla I\|$ is perturbed by an amount $\{\beta : 0 < \beta \leq 1\}$ and expressed as:

$$\|\nabla I\| = \sqrt{\|\nabla I\|^2 + \beta} \tag{5.3}$$

so that solution of the perturbed problem will converge to a solution of the original expression as $\beta \to 0$ [125]. After perturbation the solution $I_{\sigma\sigma}$ becomes:

$$I_{\sigma\sigma} = \frac{(I_x(x,y)^2 + \beta) I_{yy}(x,y) - 2 I_x(x,y) I_y(x,y) I_{xy}(x,y)}{(I_x(x,y)^2 + I_y(x,y)^2 + \beta)^{3/2}} + \frac{I_{xx}(x,y)(I_y(x,y) + \beta)^2}{(I_x(x,y)^2 + I_y(x,y)^2 + \beta)^{3/2}} \tag{5.4}$$

5.3 Gradient Descent Minimization

Based on the theoretical formulation of the proposed algorithm, the variance step was fixed at 0.1 in the gradient descent minimization scheme:

$$I_{\sigma+1} = I_\sigma - \eta_\sigma \nabla (I_\sigma) \tag{5.5}$$

The notations $I_{\sigma+1}$, I_σ, η_σ and $\nabla(I_\sigma) = I_{\sigma\sigma}$ are the currently observed image, initial image, variance step and gradient of the initial image. Minimization of image gradient is subject to a mathematical model describing relationship between Markov random field energy and noise level. At each iteration Markov random field energy expressed by total clique potential is computed according to the formulation in [117]. The stopping criterion E_s is fixed at $E_s = K_{bg}U_{bg}$ and $E_s = K_{fg}U_{fg}$ for background and foreground modes respectively, where K_{bg} and K_{fg} are arbitrary constants. These constants are multiplication factors, derived from experimental observations and determines number of iterations before energy E_s is attained in the minimization process. The images for the experiments were with background and K_{bg} was set at 2.5.

5.4 Comparative Performance Evaluation

The proposed algorithm was compared alongside three state-of-the-art algorithms. On the evaluation figure the proposed approach is denoted **CPFM**. The state-of-the-art algorithms are the nonlocal means algorithm proposed in [46], denoted **NLOC**, anisotropic diffusion approach developed in [126], denoted **DIFF** and implemented by D. S Lopes in `(http://web.ist.utl.pt/daniel.s.lopes/software)` and the wavelet technique as detailed in [10], denoted **WALT**. The comparative performance evaluation experiment was biased in favor of **NLOC** because noise level, an input parameter for its optimal performance was assumed to be accurately estimated. Evaluation was based on the following five criteria: visual quality, signal-to-noise-ratio, structural similarity index, mean square error and root mean square error.

Visual Quality

Images in Fig. 5.1a and Fig. 5.2a are clean MRI slice images from NeuroRx and ADNI, respectively. The degraded version of the image at noise levels of $\sigma = 30$ is displayed in Fig. 5.1b and Fig. 5.2b. Visual quality of the denoised version from the proposed algorithm and other algorithms can be said to be comparable to each other. Based on visual cues it can be said that the proposed algorithm and nonlocal means **NLOC** are competing for superiority. At first glance, **NLOC** can be said to be the best based on global view of the images. There is no doubt that **NLOC** performed excellently in removing noise from the background pixels, but the background region of the images are useless in medical diagnostics. A cursory examination of foreground regions in all the images shows that the proposed algorithm **CPFM** was the best in preserving original structures of the images. The white and grey matter structures in the MRI images denoised by the proposed algorithm are superior in visual quality because they are much more similar to the original images than images denoised by **NLOC** which exhibit some degree of blurring.

Signal-to-Noise Ratio

The proposed algorithm is comparable to other algorithms in terms of signal-to-noise ratio as shown in Fig. 5.3a and Fig. 5.4a. In both cases the plot identifying the proposed algorithm is raised above other algorithms but trailing behind **NLOC**. As shown in the figure, for noise level in the range $(0 \leq \sigma \leq 75)$, nonlocal means technique can be distinguished as a superior algorithm.

Structural Similarity Index

For evaluation based on the structural similarity index (Fig. 5.3b) and (Fig. 5.4b) all the algorithms are comparable. However, the best performance indicators of 0.57 and 0.55 were recorded by **NLOC**. The proposed algorithm was trailing behind **NLOC** with 0.47 and 0.51 and ahead of other state-of-the-art algorithms **DIFF** and **WALT**.

Mean Square Error

The plots in Fig. 5.3c and Fig. 5.4c on mean square error show that the plot identifying the proposed algorithm is lower than other algorithms but higher than **NLOC**, an indication that the proposal has the second best performance indicator after **NLOC** for noise level in the range $0 \leq \sigma \leq 75$.

Root Mean Square Error

In Fig. 5.3d and Fig. 5.4d the proposal recorded root mean square error of 120 in both cases, and was trailing behind nonlocal means which recorded the best performance indicator of 50 and 60, respectively.

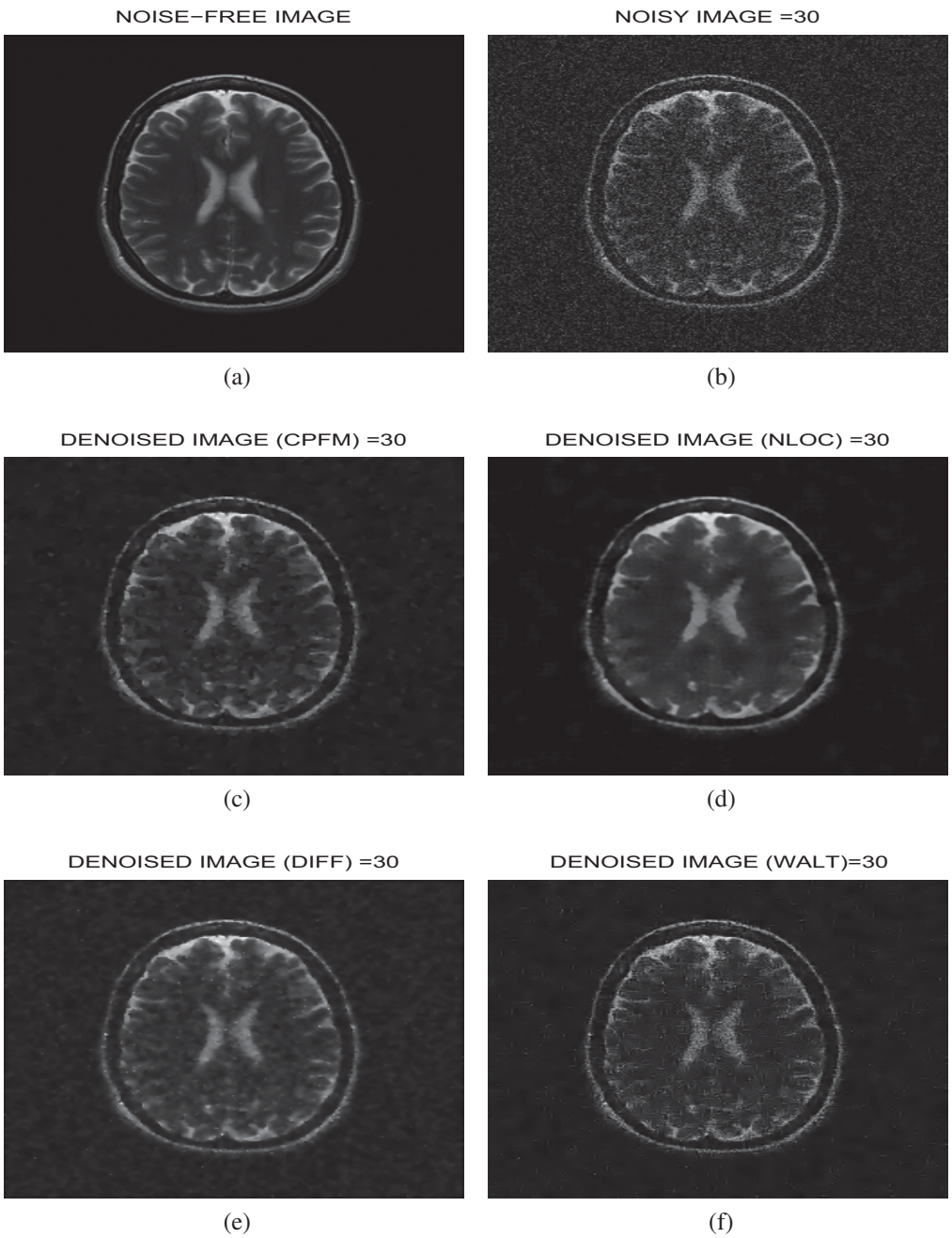


Figure 5.1: Comparative performance evaluation using MRI slice image from NeuroRx Research Inc. (a) The original and clean MRI (b) The clean image degraded by noise level of $\sigma = 30$ (c) Denoised version of (b) by **CPFM** (d) Denoised version of (b) by **NLOC** (e) Denoised version of (b) by **DIFF** (f) Denoised version of (b) by **WALT**.

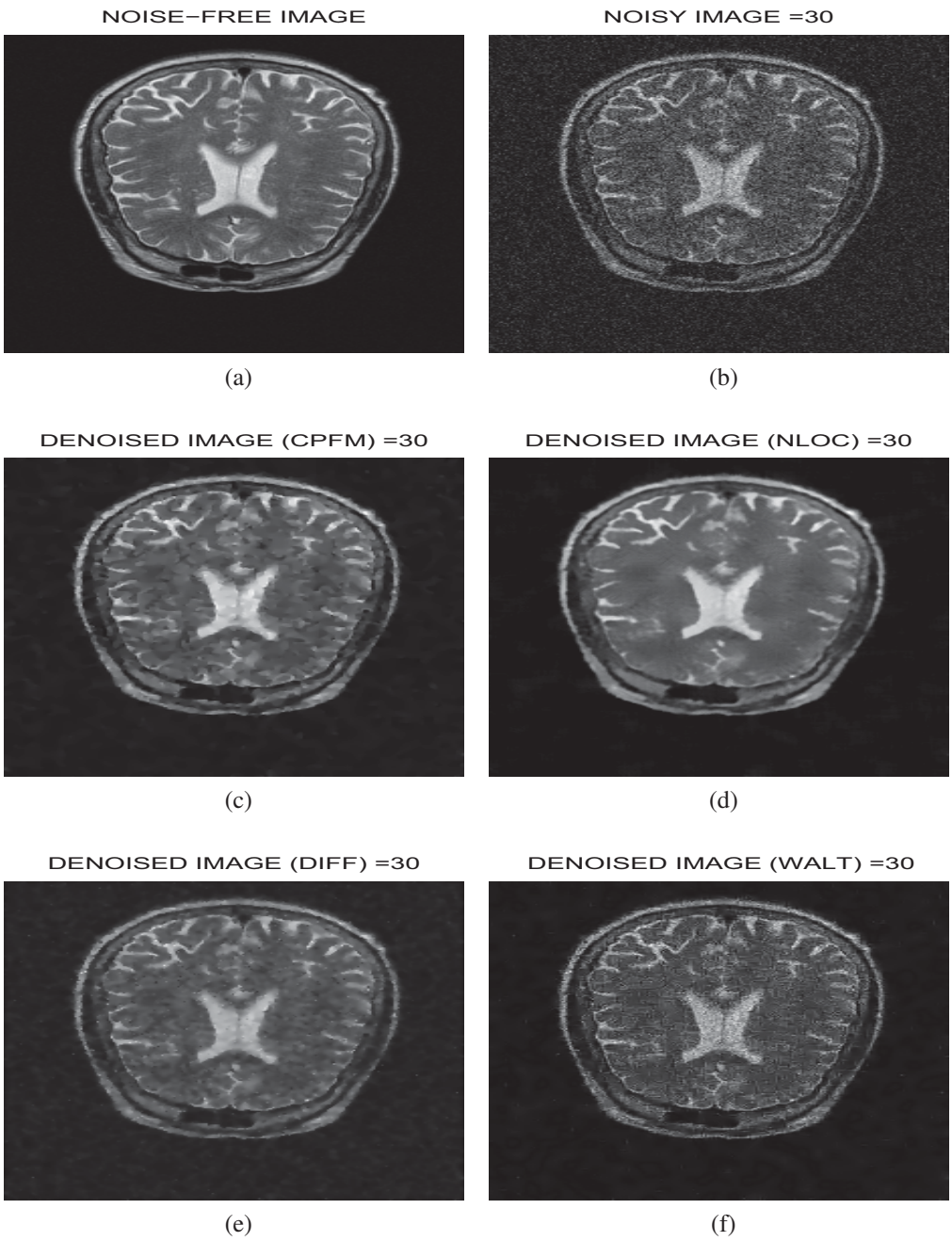


Figure 5.2: Comparative performance evaluation using MRI slice image from Alzheimer's Disease Neuroimaging Initiative. (b) The clean image degraded by noise level of $\sigma = 30$ (c) Denoised version of (b) by **CPFM** (d) Denoised version of (b) by **NLOC** (e) Denoised version of (b) by **DIFF** (f) Denoised version of (b) by **WALT**.

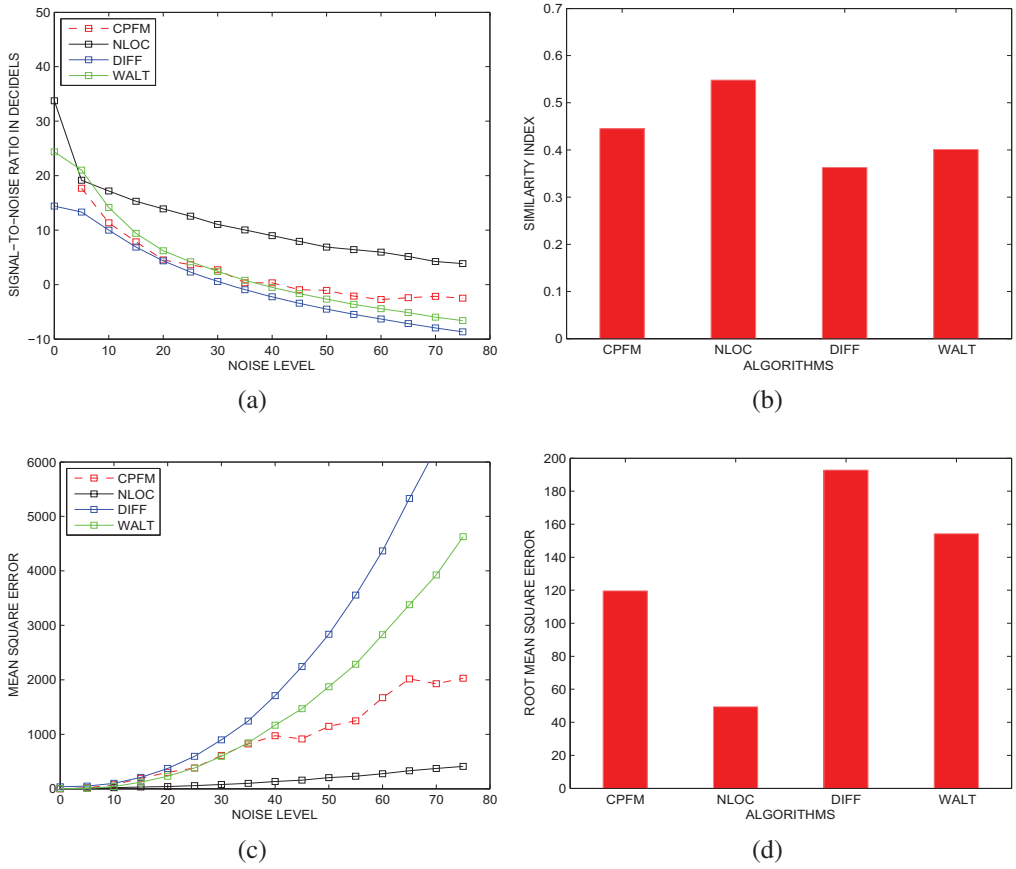


Figure 5.3: Comparative performance evaluation over the range of noise levels $0 \leq \sigma \leq 75$ using MRI slice image from NeuroRx Research Inc. (a) Signal-to-noise-ratio (b) Structural similarity index (c) Mean square error (d) Root mean square error.

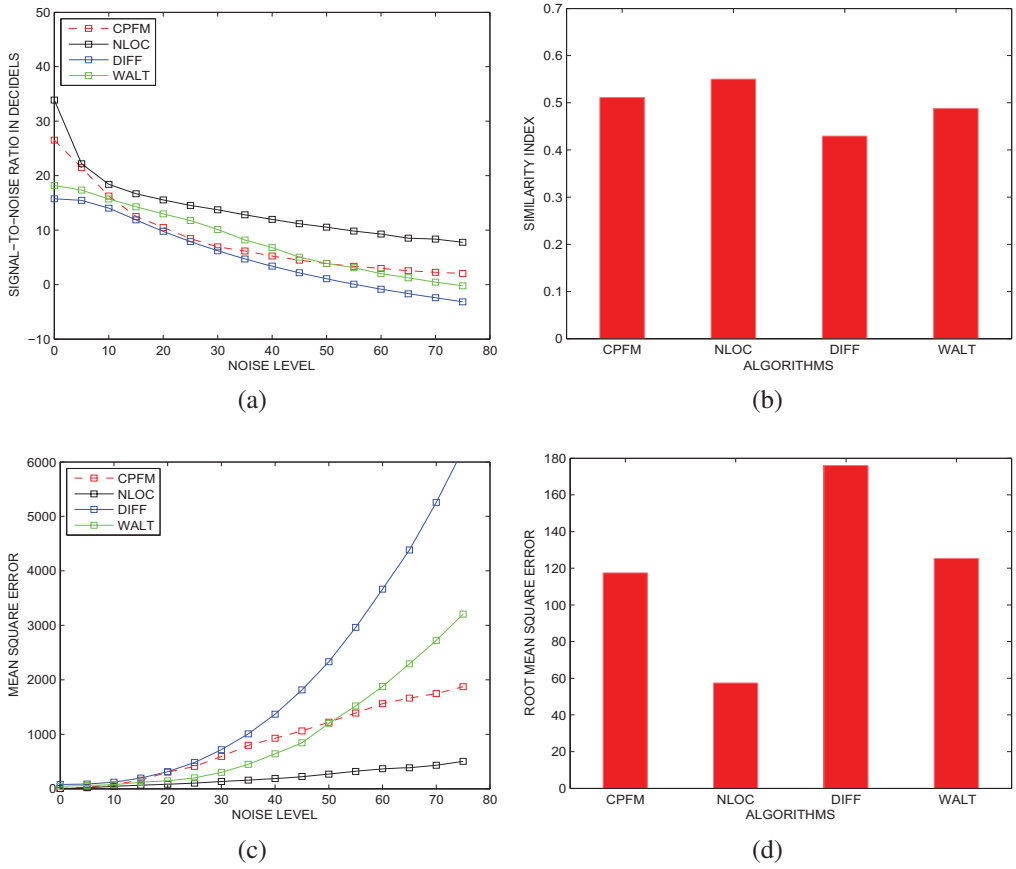


Figure 5.4: Comparative performance evaluation over the range of noise levels $(0 \leq \sigma \leq 75)$ using MRI slice image from Alzheimer's Disease Neuroimaging Initiative. (a) Signal-to-noise-ratio (b) Structural similarity index (c) Mean square error (d) Root mean square error.

Chapter 6

A General Framework for Denoising

This chapter proposes two approaches to generalize the application of Markov random field model to noise reduction in images [121]. A first approach is the concept of the Variational-Bayesian cycle. It is a theoretical formulation on how to *'navigate'* from ROF total variation to a classical Markov random field model and *vice-versa*. Based on the concept of the VB cycle, it is shown that the variational and Bayesian techniques are equivalent and noise variance is the optimal regularization parameter in TV denoising of MRI images of the brain [127]. The second approach describes the design of an algorithm which can be configured for Bayesian and variational denoising modes by assigning scalar and vector values to the smoothness energy, respectively. Within the variational mode, the choice of norm adapts images for either the total variation or the Tikhonov technique.

6.1 The Variational-Bayesian Cycle

The focus of researchers on the computational speed of TV in different applications has led to generalizations and variations of the original TV denoising problem formulated by ROF [62]. In this paper the original TV problem formulation expressed as a convex functional in [122] was adopted for the proposed method:

$$\hat{I}_c = \min_{I_d} \quad \sum_{i=1}^{M}\sum_{j=1}^{N} \|(\nabla \mathbf{I_d})_{(i,j)}\| + \frac{1}{2\lambda}\|I_c - I_d\|^2 \tag{6.1}$$

The observed image I_d is of size MXN and the other notations have same meaning as in Eq. 3.16. Natural logarithmic transformation on Eq. 6.1 maintains the equality of the right hand side (RHS)

and left hand side (LHS) terms of the equation and also retains the convexity of the functional:

$$\ln\left[\exp-(\hat{I}_c)\right] = \min_{I_d} \quad \ln\left[\exp-\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\|(\nabla\mathbf{I_d})_{(i,j)}\| + \frac{1}{2\lambda}(I_c - I_d)^2\right)\right] \tag{6.2}$$

Now to separate the variables on RHS of Eq. 6.2 into product of two exponential functions:

$$\ln\left[\exp-(\hat{I}_c)\right] = \min_{I_d} \quad \ln\left[\exp-\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\|(\nabla\mathbf{I_d})_{(i,j)}\|\right)\exp-\left(\frac{1}{2\lambda}(I_c - I_d)^2\right)\right] \tag{6.3}$$

With reference to previous work [117] the first term on the RHS of Eq. 6.3 is the single layer Markov random field energy $U(f)$ expressed as a function of the image pixel configuration f:

$$U(f) = \sum_{i=1}^{M}\sum_{j=1}^{N}(\|(\nabla I_d(f))_{(i,j)}\|) \tag{6.4}$$

Insert this term into Eq. 6.3:

$$\ln\left[\exp-(\hat{I}_c)\right] = \min_{I_d} \quad \ln\left[\exp-(U(f))\exp-\left(\frac{1}{2\lambda}(I_c - I_d)^2\right)\right] \tag{6.5}$$

Expressing RHS and LHS terms of Eq. 6.5 as strictly exponential functions maintains equality of both sides of the equation but the functional changes from a convex functional to a concave functional. Thus the optimization criteria of the functional changes from minimization to maximization:

$$\left[\exp-(\hat{I}_c)\right] = \max_{I_d} \quad \left[\exp-(U(f))\exp-\left(\frac{1}{2\lambda}(I_c - I_d)^2\right)\right] \tag{6.6}$$

In the classical Markov random field model [111] the Gibbs distribution $P(f)$, the probability distribution of each possible configuration of the prior I_c is proportional to the first exponential term on the RHS of Eq. 6.6:

$$P(f) = P(I_c) \propto [\exp-(U(f))] \tag{6.7}$$

The Gaussian distribution P_N, the probability distribution assumed for the observed image I_d given any realization of I_c is proportional to the second exponential term on the RHS of Eq. 6.6:

$$P_N = P(I_d|I_c) \propto \left[\exp-\left(\frac{(I_c - I_d(f))^2}{2\lambda}\right)\right] \tag{6.8}$$

Thus:

$$\exp-\left[\frac{(I_c - I_d(f))^2}{2\lambda}\right] \equiv \exp-\left[\frac{(\mu - \mathbf{X})^2}{2\sigma^2}\right] \tag{6.9}$$

where the mean μ of the Gaussian distribution is the underlying image I_c, the observed image I_d is the variable $\mathbf{X}$ of the Gaussian distribution. Thus, the regularization parameter λ is equal to the noise variance σ^2 of the image:

$$\lambda = \sigma^2 \tag{6.10}$$

Insert the probability notations on the RHS of Eq. 6.7 and Eq. 6.8 into RHS of Eq. 6.6 This gives the Bayes posterior probability $P(I_c|I_d)$ formula [114]:

$$P(I_c|I_d) \propto [P(I_d|I_c)P(I_c)] \tag{6.11}$$

Each possible configuration $f \in F$, where F is a discrete set of random variables, is independent and identically distributed. Thus Eq. 6.11 can be expressed as

$$P(I_c|I_d) \propto \left\{ \left[\prod_{f \in F} \exp -(U(f)) \right] \left[\prod_{f \in F} \exp - \left(\frac{1}{2\sigma^2} (I_c - I_d(f))^2 \right) \right] \right\} \tag{6.12}$$

and arrive at the Bayesian formulation of the classical Markov random field model for image restoration introduced by Geman and Geman [60]:

$$\begin{aligned} \hat{I_c} &= \arg\max_{I_d} \quad P(I_c|I_d) \\ &= \arg\max_{I_d} \quad [P(I_d|I_c)P(I_c)] \\ &= \arg\max_{I_d} \quad \left\{ \left[\prod_{f \in F} \exp -(U(f)) \right] \left[\prod_{f \in F} \exp - \left(\frac{1}{2\sigma^2} (I_c - I_d(f))^2 \right) \right] \right\} \end{aligned} \tag{6.13}$$

This formulation estimates the denoised image $\hat{I_c}$ from the image configuration which maximizes $P(I_c|I_d)$ in what is referred to as maximum *a posteriori* probability (MAP). The probability distributions are functions of the image pixel configurations. The configurations are determined by the level of noise and the level of noise is a function of the Markov random field energy.

MAP can be reformulated by adopting a two-step process. The first step is to regard MAP as minimization of the negative exponential terms in Eq. 6.13. The second step is natural logarithm transformation of the resulting exponential function:

$$\hat{I_c} = \arg\min_{I_d} \quad \left\{ U(f) + \frac{(I_c - I_d(f))^2}{2\sigma^2} \right\} \tag{6.14}$$

Replacing the image energy in the first term of Eq. 6.14 with the term on the RHS of Eq. 6.4 and

substituting the expression for σ in Eq. 6.10 into the second term of Eq. 6.14:

$$\hat{I}_c = \arg\min_{I_d} \left[\sum_i \sum_j \|(\nabla I_d(f))_{(i,j)}\| + \frac{1}{2\lambda} \|I_c - I_d(f)\|^2 \right] \tag{6.15}$$

which is a return back to variational mode of denoising expressed in Eq. 6.1, thus completing a full circular path to confirm the equivalence of variational and Bayesian frameworks. The circular path is referred to as the Variational-Bayesian cycle. Graphical expression of the VB cycle is shown in Fig. 6.1.

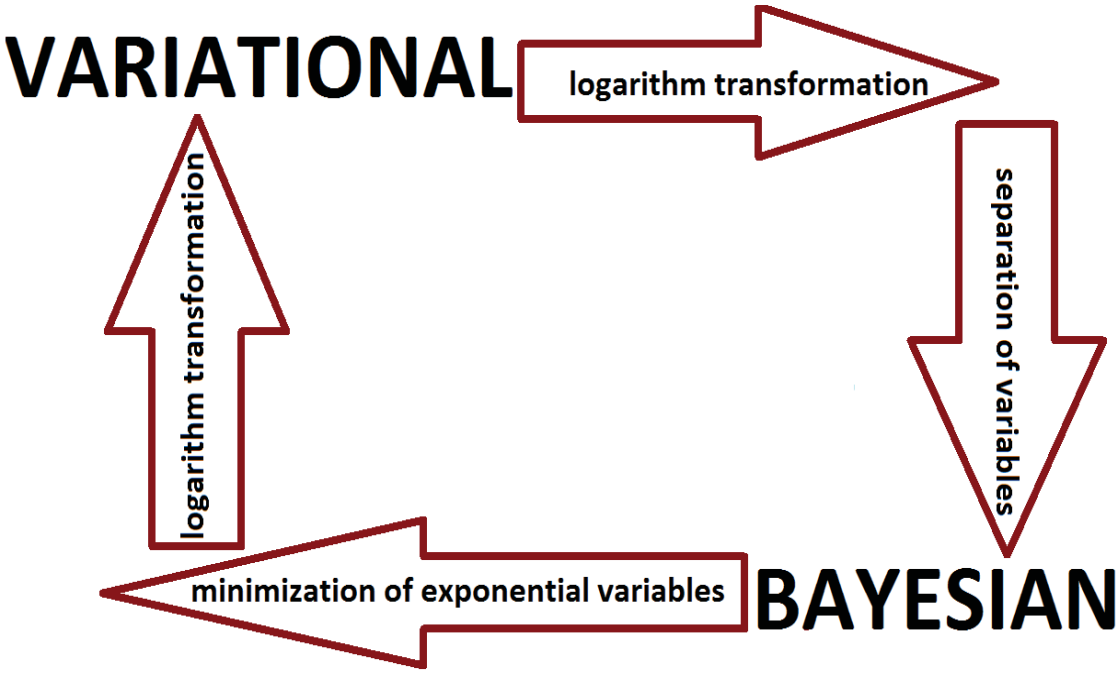


Figure 6.1: The Variational-Bayesian (VB) cycle. Logarithmic transformation of the variational problem formulation is followed by the separation of exponential variables, and this results in Bayesian problem formulation. A return to the variational problem formulation is by minimization of the exponential variables in the Bayesian problem formulation followed by a logarithmic transformation.

6.2 Optimal Selection of Regularization Parameter

This section demonstrates that the noise variance derived in Eq. 6.10 is the optimal regularization parameter for a given image. Given a brain MRI image, its Markov random field energy, expressed by total clique potential energy, is computed. The value of this energy is used to estimate the variance of the image from the mathematical model that describes the relationship between Markov random field energy and noise variance for brain MRI images [117]. The variance of the

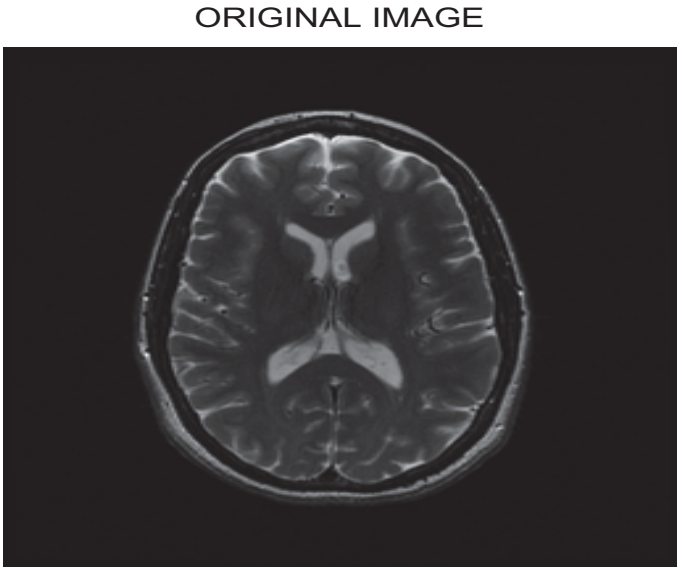
image is further used to estimate, at no computational cost, the regularization parameter before commencement of the TV denoising process. The estimated regularization parameter is scaled to seven different values from 0.1λ to 1.9λ in steps of 0.3 and the denoised images associated with the different values are recorded.

Fig. 6.2 displays images of a T2-weighted MRI slice from NeuroRx. The slice is indexed as slice number 32 in a single subject MRI data consisting of 45 slices. The original, degraded and denoised images are in Fig. 6.2a, Fig. 6.2b, and Fig. 6.2c, respectively. Optimality of the different scales of the computed regularization parameter in terms of mean square error (MSE) convergence is shown in Fig. 6.3

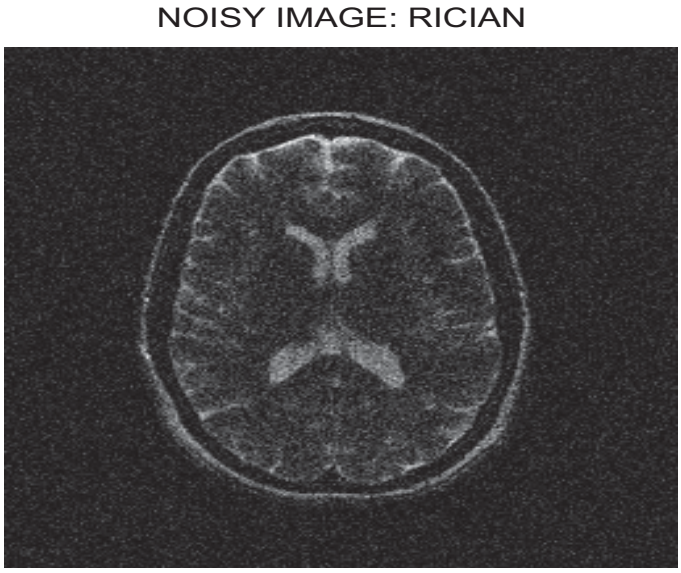
In Fig. 6.4 are images of a T2-weighted MRI slice from ADNI. The slice is indexed as slice number 28 in a single subject MRI data consisting of 42 slices. The original, degraded and denoised images are in Fig. 6.4a, Fig. 6.4b and Fig. 6.4c, respectively. Optimality of the different scaled values of the computed regularization parameter in terms of MSE convergence is shown in Fig. 6.5.

6.2.1 Optimality Test

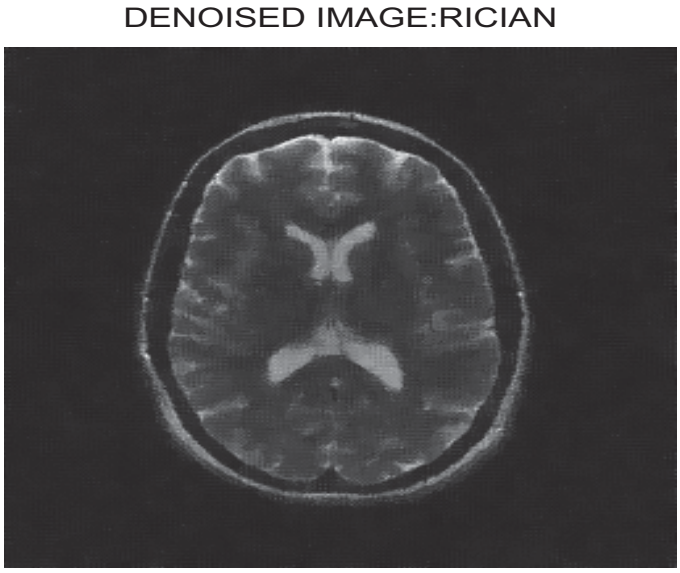
Optimality test, in terms of MSE using regularization parameters generated from the set $\{0.1\lambda_{Op} : 0.3 : 2\lambda_{Op}\}$, are the plots shown in Fig. 6.3 and Fig. 6.5. A cursory look at the plot indicates that three regularization parameter values $0.7\lambda_{Op}$ (blue colored solid line), λ_{Op} (red colored solid line with circle) and $1.3\lambda_{Op}$ (black colored dash line) are candidates for the optimal value. The plots identifying $1.3\lambda_{Op}$ in Fig. 6.3 and Fig. 6.5 shows a 23 percent and a 27 percent decrease, respectively, in MSE after about 10 iterations but lost its steam afterwards by reversing its earlier gains to 20 percent and 25 percent, respectively, after 100 iterations. The plots identifying the parameter $0.7\lambda_{Op}$ recorded a 25 percent (see Fig. 6.3) and a 30 percent (see Fig. 6.5) decrease in MSE after 40 iterations. The plot identifying λ_{Op} reduced the MSE of the degraded image by 25 percent (see Fig. 6.3) and by about 30 percent (see Fig. 6.5) in 20 iterations and it maintained this same level of performance for up to 100 iterations. 20 iterations is half the number of iterations it takes the parameter $0.7\lambda_{Op}$ to attain same level of performance. Thus, it is reasonable to conclude that the plot identifying λ_{Op} is the optimal value of λ.



(a)



(b)



(c)

Figure 6.2: (a) A T2-weighted MRI slice image from NeuroRx Research in its (a) original state of acquisition (b) degraded state by noise level of $\sigma = 25$ (c) denoised version using the TV algorithm with $\lambda = 25$.

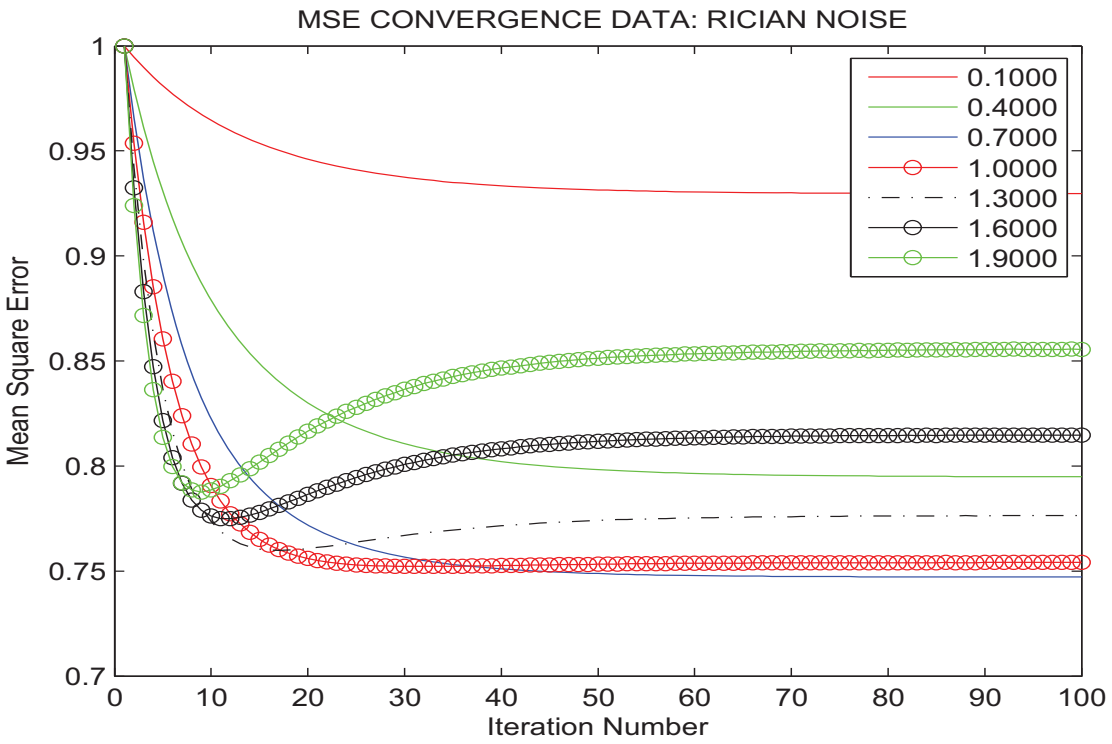
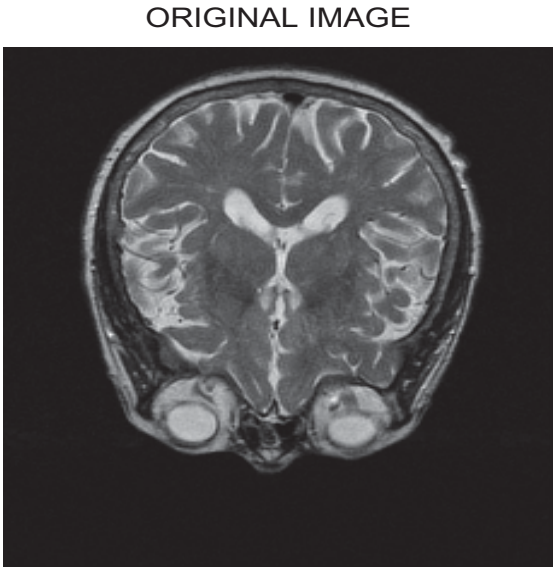
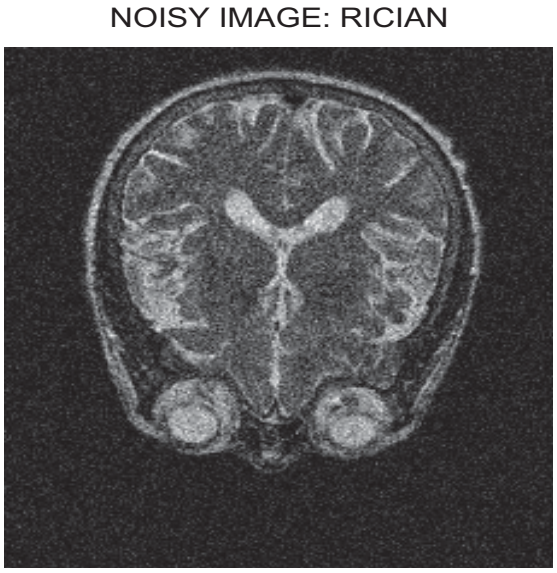


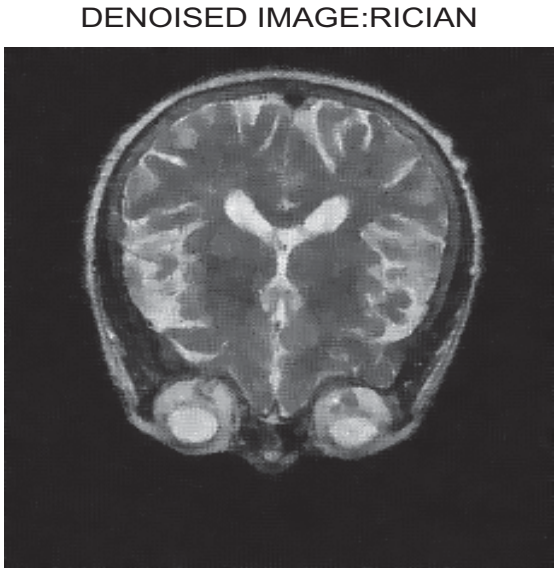
Figure 6.3: Test of optimality of computed regularization parameter. The MRI slice image from NeuroRx shown in Fig.6.2b was denoised using TV algorithm with the computed optimal regularization parameter scaled from 0.1 to 2 at interval of 0.3.



(a)



(b)



(c)

Figure 6.4: (a) A T2-weighted MRI slice image from ADNI in its (a) original state of acquisition (b) degraded state by noise level of $\sigma = 25$ (c) denoised version using TV algorithm with $\lambda = 25$.

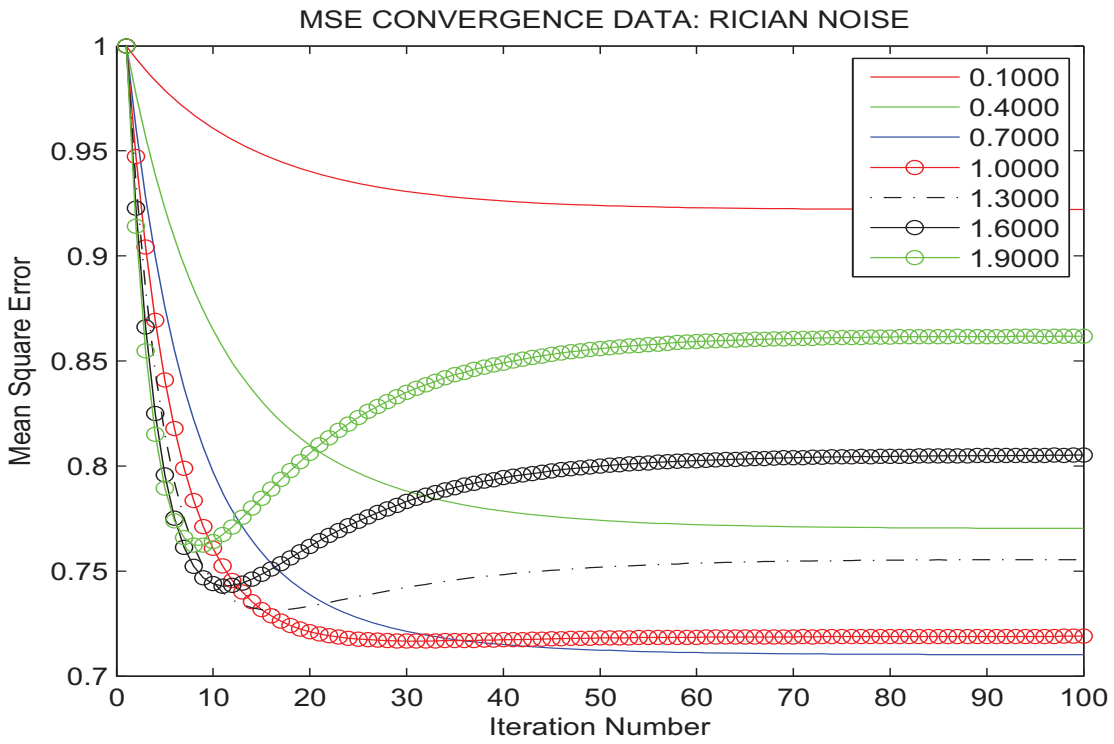


Figure 6.5: Test of optimality of computed regularization parameter. The MRI slice image from ADNI shown in Fig.6.4b was denoised using TV algorithm with the computed optimal regularization parameter scaled from 0.1 to 2 at interval of 0.3.

6.3 Algorithm Design

The design of the proposed general framework is illustrated by the flow chart in Fig. 6.6. The system has three input parameters. The first is the number of layers $(\mathbf{SL}, \mathbf{DL})$ representing the number of images under consideration. The second is the type of value $(\mathbf{V}, \mathbf{S})$ assigned to ξ_r and ξ_p for conformity and violations of the smoothness constraints. The third parameter is the choice of norm (**q=1,2**) for computing magnitude of the image gradient. For a user's selection of single layer **SL**, the denoising system does not make reference to a prior image model I_c but a vector valued kernel of Eq. 3.10 is used for computing the gradient of the observed image I_d according to Eq. 3.14. For specific data such as the brain MRI having pixels in the background, the energy minimization is controlled by the mathematical model stated in Eq. 4.3. At each iteration the Markov random field energy is computed and the stopping criterion is fixed at $(E_s = K_b U_b = -1.5)$ based on knowledge of mathematical function describing relationship between MRF energy and noise level [117].

The choice of double layer **DL** introduces a prior image as the reference and requires the user to make an additional choice of either scalar or vector valued smoothness constraint. With the choice of scalar value **S** the classical Markov random field algorithm is deployed by the system to operate on the image and the energy of the image is computed according to Eq. 3.13. For the choice of vector valued smoothness **V** the system prompts the user to a choice of either L_1 or L_2 according to the selection of $\mathbf{q} = 1$ or $\mathbf{q} = 2$, respectively. Either choice deploys the total variation or Tikhonov technique to operate on the image and the energy of the image is computed according to Eq. 3.14.

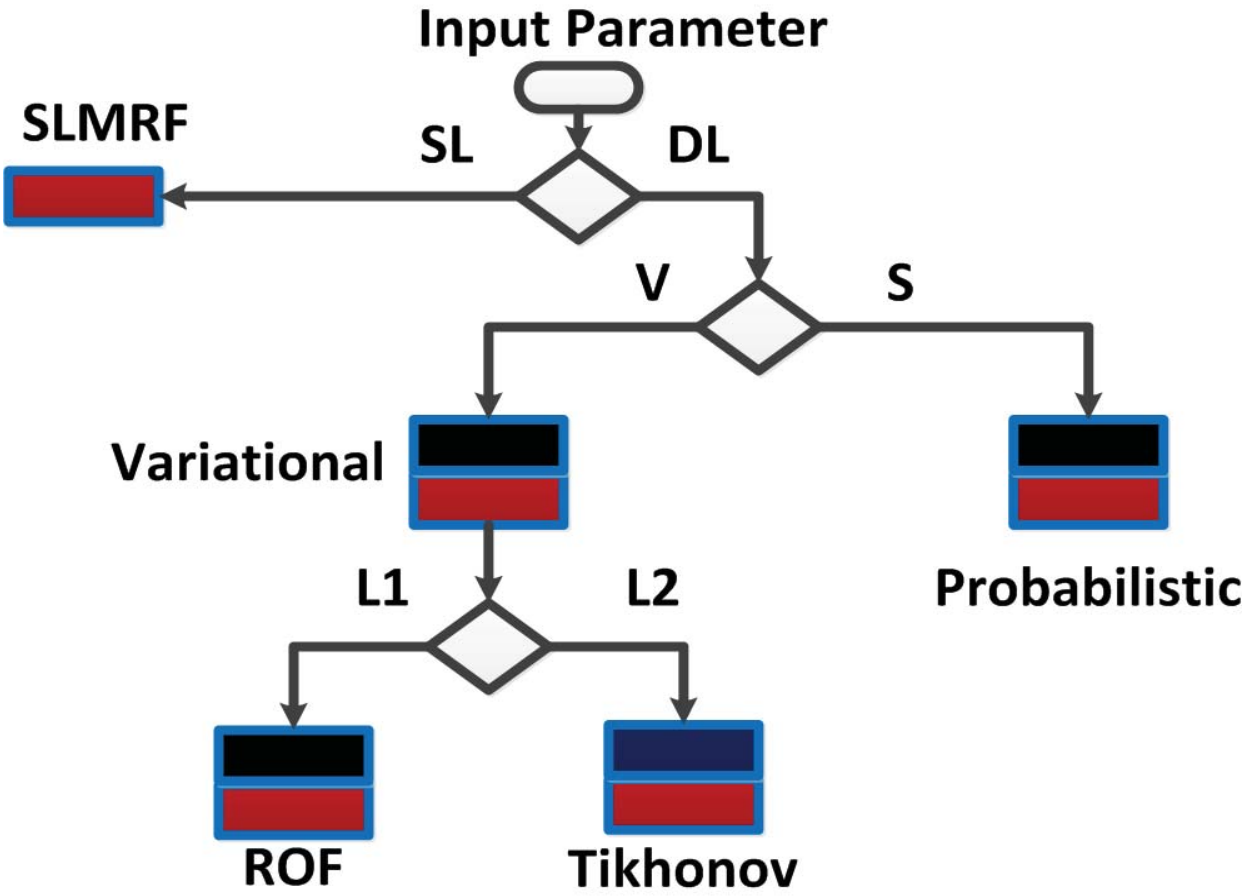


Figure 6.6: From the user choice of single layer (**SL**) there are two paths from the proposed single layer Markov random field (SLMRF) model having only the observed image (red color) to double layer (**DL**) Markov random field model having both observed and prior model images (black color). Scalar (**S**)and vector (**V**) values assigned for reward and penalty for conformity and violations of the smoothness constraints configures the image for double layer probabilistic and variational approaches respectively. Choice of L_1 and L_2 norm for the image gradient leads to Tikhonov regularization and Rudin, Osher and Fatemi (ROF) total variation techniques, respectively.

Chapter 7

Intensity Inhomogeneity Correction

A bias field correction algorithm is proposed in this chapter [128]. It is a new model which is a true reflection of the presence of a bias field in MRI imaging process. The new model does not require an estimation of a bias field before determination of the restored or corrected image. The proposal is a three-step bias field correction strategy:

(1) Create regions of interest (ROI) by clustering. Number of ROI should be cleverly chosen to equal number of anatomical structures under investigation.

(2) Compensate for vagueness of intensity level attributes by seeking extra information from an anatomical structural map.

(3) Detect voxel locations where the bias field is considered significant using an anatomic structural map.

The proposed algorithm was tested on only real MRI images. Image simulators were not considered as test data to validate the proposed algorithm because they lack the natural anatomical variability and image acquisition artifacts of real images [13]. Fifteen T1-weighted brain MRI images with different severity levels of bias fields were sourced from NeuroRx Research Inc.

The design of the algorithm is described in next section, followed by display of the experimental results. The performance of the algorithm was evaluated qualitatively using visual inspection of corrected images. Visual quality assessment, though subjective, was preferred because it is the only performance evaluation criterion that validates all objective performance criteria. Quantitative comparative performance evaluation was avoided because of lack of publicly available ground truth data. Moreover, credible implementation methods and optimal operating parameters of other algorithms is challenging because it requires interaction with the original authors [13]. The chapter ends with a discussion of experimental results

7.1 Algorithm Design

Implementation of the proposal is explained by the flow chart in Fig. 7.1. The test image is at top of the figure. Red-colored numbers in the figure depict each step for the implementation of the algorithm. The first step thresholds the test image using the Otsu method [129]. The Otsu method finds the threshold that minimizes intra-class variations within the image. The Otsu method is preferred for thresholding because in many applications, Otsu method has proved to be simple and effective [130]. Noise and voxels corrupted by a bias field are eliminated. Only very strong edges representing image details are retained. Output of the Otsu method is one of two components needed to create a template of the image. The second step extracts foreground voxels of the test image. In the third step the test image is fed to a multi-scale canny edge detector. The canny edge detector [131] is the preferred edge detector because of its optimal performance in terms of robustness to noise, minimal response and good location of edges [132]. The multi-scale edge detector outputs two images. The first image is an edge image at the minimum edge scale. The other is an edge image at the maximum edge scale. In fourth step, the edge image at maximum scale is combined with the Otsu image to produce an anatomical structural map of the test image. In the fifth step, regions of interest (ROI) in the image are generated by the k-means algorithm. The number of regions of interest is chosen to equal the number of tissue classes under investigation. The choice of k-means instead of a more robust algorithm was motivated by its simplicity, an important requirement at the preliminary steps of the proposed algorithm. The sixth step tests each region of interest for outliers using entropy and spatial information from the structural map according to Eq. 3.35 and Eq. 3.36. Outliers are assigned to their appropriate tissue class j in the seventh step to produce new sets of voxels for each tissue class. The new sets of voxels $x_j \in [p_{1j} \quad p_{2j}]$ are rescaled by mapping to $x_j^{'} \in [s_{1j} \quad s_{2j}]$ [133].

$$x_j^{'} = s_{1j} + \left(\frac{x_j - p_{1j}}{p_{2j} - p_{1j}} \right) (s_{2j} - s_{1j}) \tag{7.1}$$

where p_{1j} and p_{2j} are minimum and maximum values of re-classified voxels, respectively, and s_{1j} and s_{2j} are minimum and maximum values of the initial set of clean voxels, respectively. The statistics of the initial set of clean voxels can be substituted with the statistics of a model for the tissue class, if available.

7.2 Experimental Results

The result of applying the proposed algorithm to different levels of severity and complexity of bias fields in brain MRI images are displayed in seven different figures, Fig. 7.2, Fig. 7.3, Fig. 7.4,

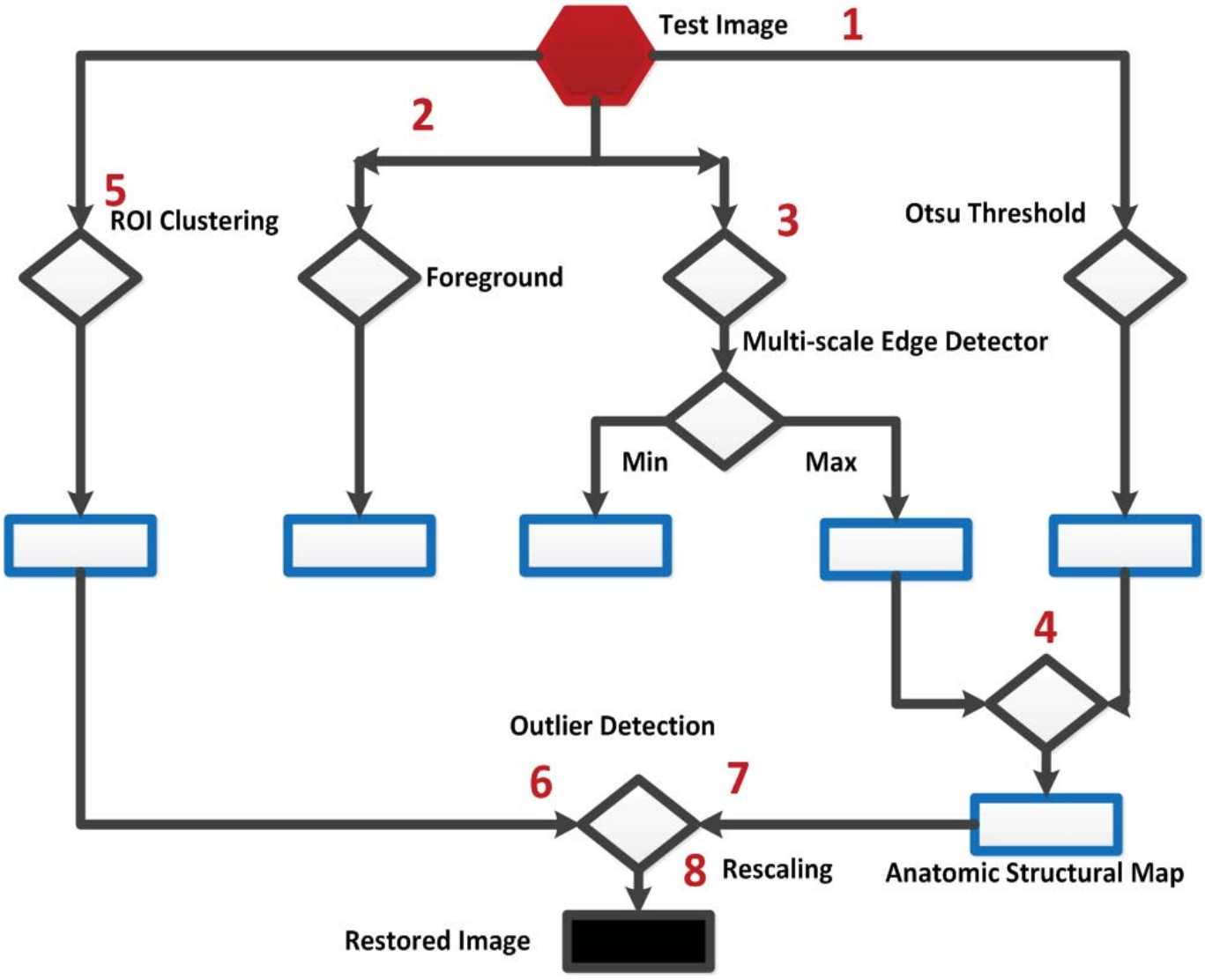


Figure 7.1: Flow chart of the bias field correction algorithm.

Fig. 7.5, Fig. 7.6, Fig. 7.7 and Fig. 7.8.

7.3 Discussion

The discussion is outlined according to the figure number of the experimental results. Discussion based on Fig. 7.2 is detailed because it explains each step in the implementation of the proposed algorithm.

7.3.1 Figure 7.2

The test image corrupted by a bias field is in Fig. 7.2a. Test image voxel values, originally acquired as int32, are converted to double precision grey level. After conversion, four duplicates of the image are made. The first duplicate, after the Otsu method of thresholding, is in Fig. 7.2b. Foreground voxels extracted from the second duplicate is in Fig. 7.2c. The third duplicate is fed to a multi-scale edge detector which outputs the edge images at minimum scale of 0.1σ and maximum scale of 0.7σ in steps of 0.1. Local entropy of both edge images are displayed in Fig. 7.2d and Fig. 7.2e. The entropy image at maximum scale in Fig. 7.2e is combined with the output of Otsu thresholding to obtain an anatomical structural map of the MRI slice. The structural map displayed in Fig. 7.2f contains accurate tissue-specific spatial and entropy information. Voxel locations with low entropy, in dark color, represent homogeneous regions of specific anatomical structures. Voxel locations with high entropy, in white color, are image details that separate homogeneous regions within and between tissue classes. Regions of interest (ROI) detected in the fourth duplicate of the test image are in Fig. 7.2g. Corrupt voxel intensity levels, known as outliers, are identified by entropy information in the tissue structural map shown in Fig. 7.2h. Outliers $G_{S,k}$ are voxel locations in one of tissue classes j in the structural map that are members of the set determined by setting entropy threshold $s_{th} = 0.3$:

$$G_{S,k} = \{\Omega_{I,j} \cap \Omega_{S,k}\} \cap \{S : s < s_{th}\} \tag{7.2}$$

High image contrast generates an edge effect [134]. In this proposal the edge effect is revealed as high entropy values recorded by the entropy filter in small homogeneous regions around edges that demarcate grey matter and white matter tissues. This problem was addressed by the area threshold $A_{th} = 100$. Area thresholding is also applied in identification of white matter regions to eliminate homogeneous regions within edges that characterize grey matter regions. The outliers are assigned to their appropriate tissue classes, rescaled to produce the restored image shown in Fig. 7.2i. Visual quality of the restored image is very high compared to the degraded image.

7.3.2 Figure 7.3

The test image in Fig. 7.3a has the same anatomical structures (except ventricles) as test image in Fig. 7.2a. The tissue structural map is shown in Fig. 7.3b. The ROI image in Fig. 7.3c shows that white matter tissues around the midline of the axial slice were misclassified as grey matter by the clustering algorithm. The restored image in Fig. 7.3d is of much higher quality than the degraded image.

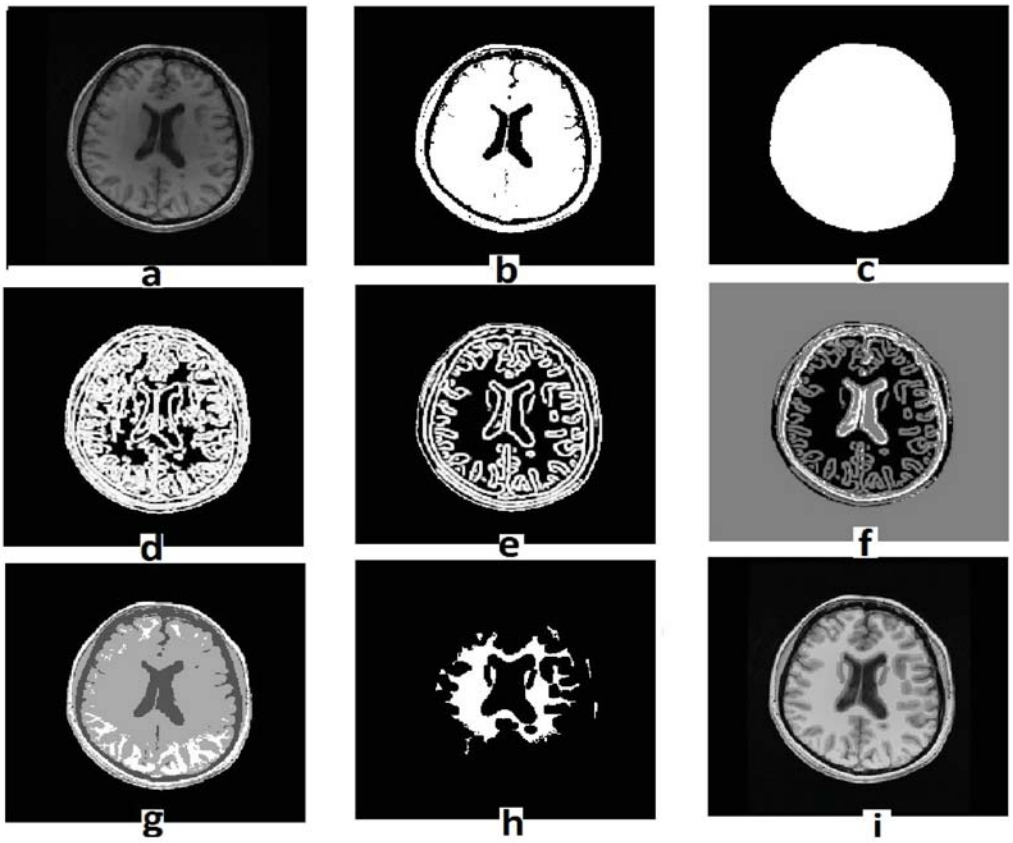


Figure 7.2: (a) Test image (b) Otsu threshold of test image (c) Foreground image (d) Entropy image at minimum scale (e) Entropy image at maximum scale (f) Anatomical structural map (g) Region of interest (h) Outliers (i) Restored image.

7.3.3 Figure 7.4

The test image in Fig. 7.4a is a typical example of a locally smooth but globally non-smooth bias field. The tissue structural map is in Fig. 7.4b. The ROI clustering algorithm correctly labeled white matter in Fig. 7.4c, though intensity levels are different within the same anatomical structure. This accurate labeling may not hold for the clustering algorithm which is more robust than the k-means algorithm. Intensity level variation was corrected by invoking the intensity rescaling function incorporated in the proposed system. The restored image is shown in Fig. 7.4d.

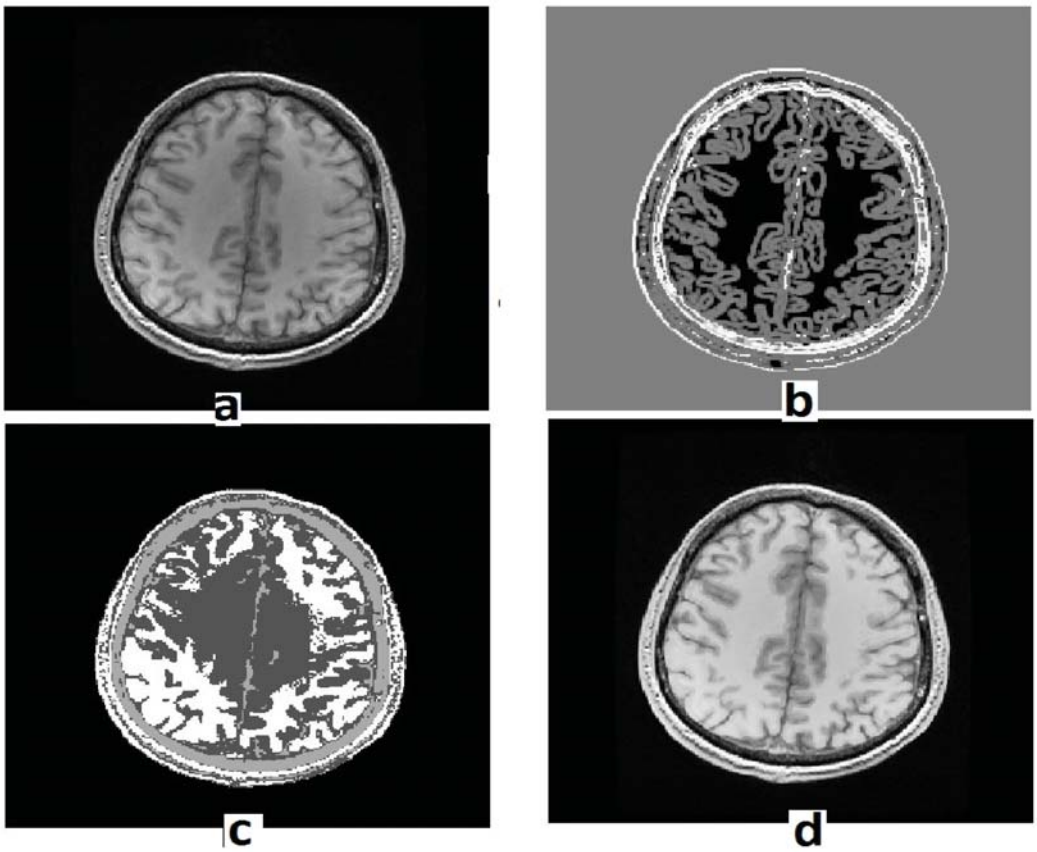


Figure 7.3: (a) Test image (b) Anatomical structural map (c) Region of interest (d) Restored image.

7.3.4 Figure 7.5

The test image in Fig. 7.5a has the same level of severity as the test image in Fig. 7.2a. The structural map in Fig. 7.5b, ROI in Fig. 7.5c and the restored image in Fig. 7.5d indicates that the proposed algorithm demonstrates the same level of performance for same level of severity of bias field.

7.3.5 Figure 7.6

The structural map in Fig. 7.6b was generated from the test image in Fig. 7.6a. Entropy and spatial information from the structural map were combined with the ROI in Fig. 7.6c to restore the image. Visual assessment of the restored image in Fig. 7.6d shows that there are much more uniform intensity levels within tissue classes than the test image.

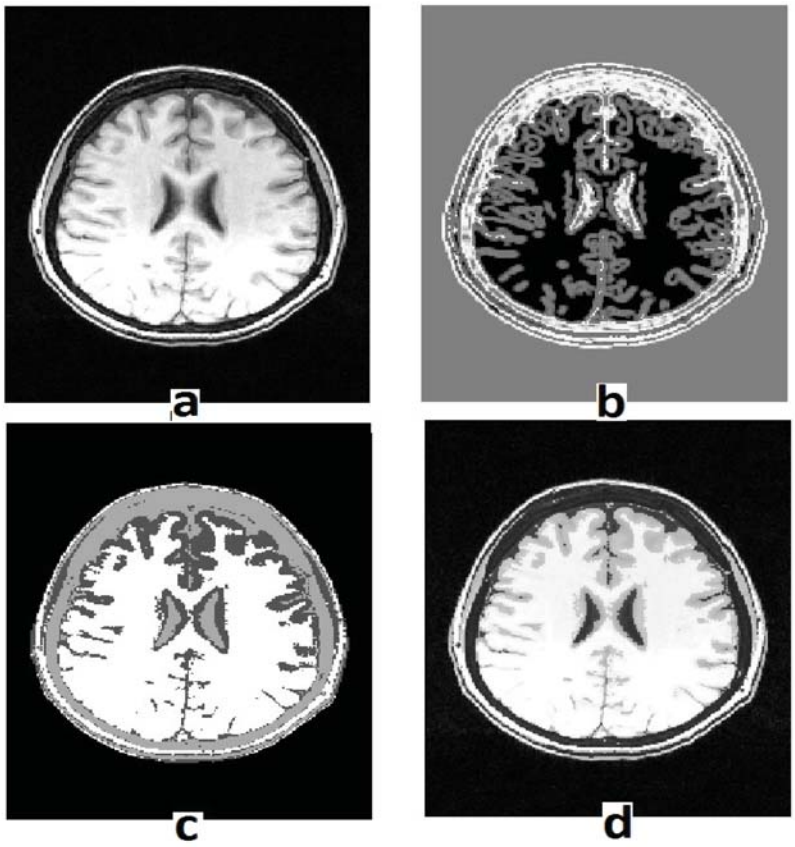


Figure 7.4: (a) Test image (b) Anatomical structural map (c) Region of interest (d) Restored image.

7.3.6 Figure 7.7

The test image in Fig. 7.7a is in same category of severity as test image in Fig. 7.4. Unlike Fig. 7.4, the ROI clustering unit in the proposed system detected the presence of a bias field. The structural map and the clustered ROI are Fig. 7.7b and Fig. 7.7c, respectively. Intensity inhomogeneity is significantly reduced in the restored image of Fig. 7.7d.

7.3.7 Figure 7.8

The test image in Fig. 7.8a is yet another example of a locally smooth but globally non-smooth bias field. The intensity level of white matter voxels towards the frontal lobe is obviously different from voxels of same anatomical structure in other regions of the MRI slice image. As depicted in Fig. 7.8c, white matter voxels towards frontal lobe were mislabeled as grey matter due to the presence of a bias field. Information from the structural map in Fig. 7.8b was combined with the

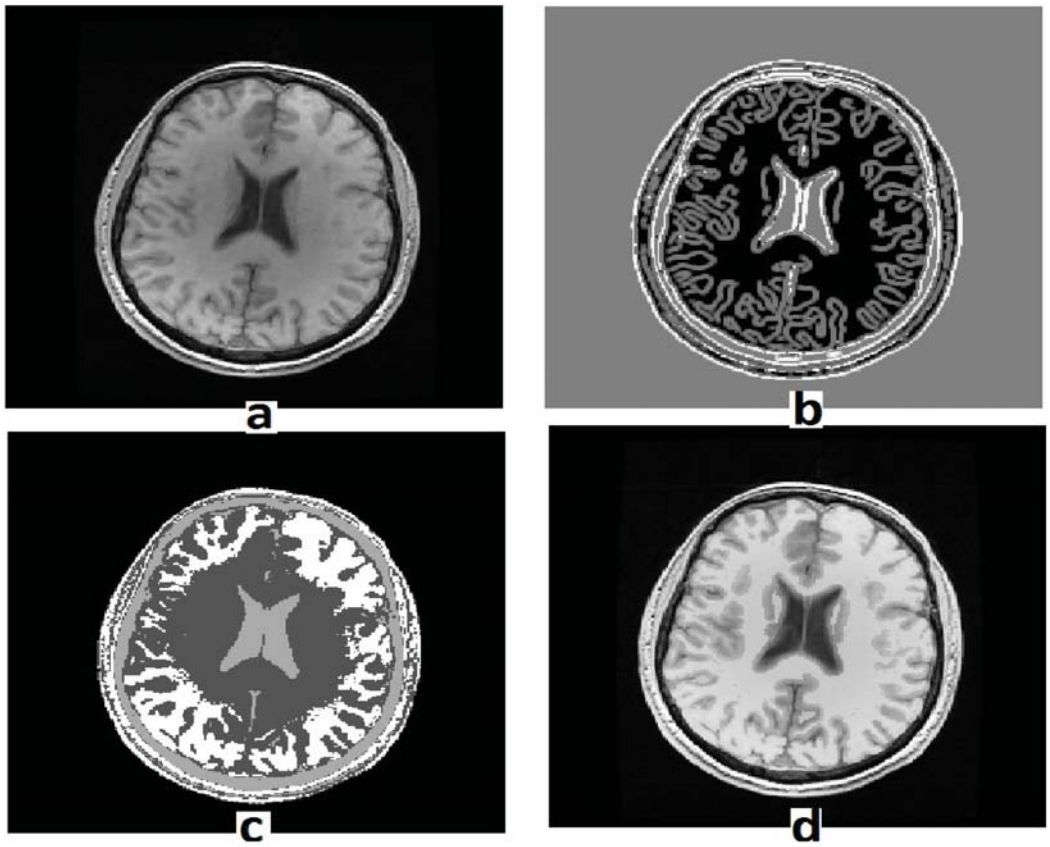


Figure 7.5: (a) Test image (b) Anatomical structural map (c) Region of interest (d) Restored image.

ROI in Fig. 7.8c to restore the image. The restored image shown in Fig. 7.8d is superior in visual quality compared to the test image.

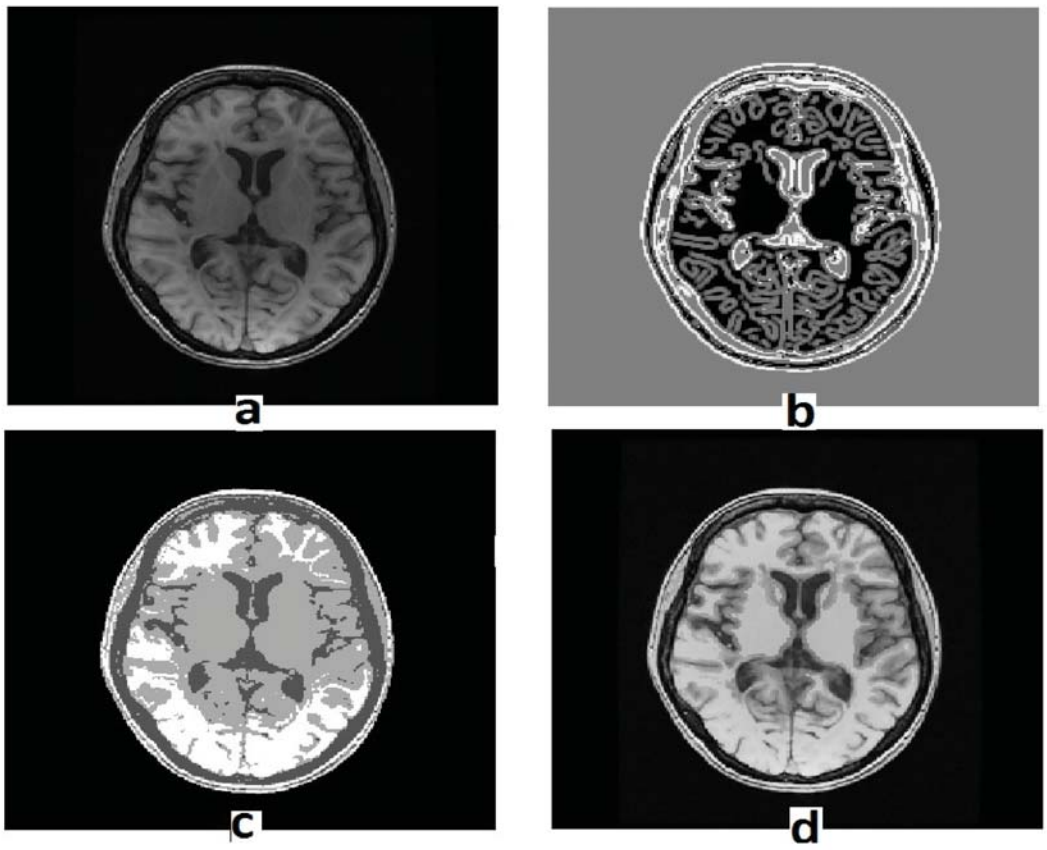


Figure 7.6: (a) Test image (b) Anatomical structural map (c) Region of interest (d) Restored image.

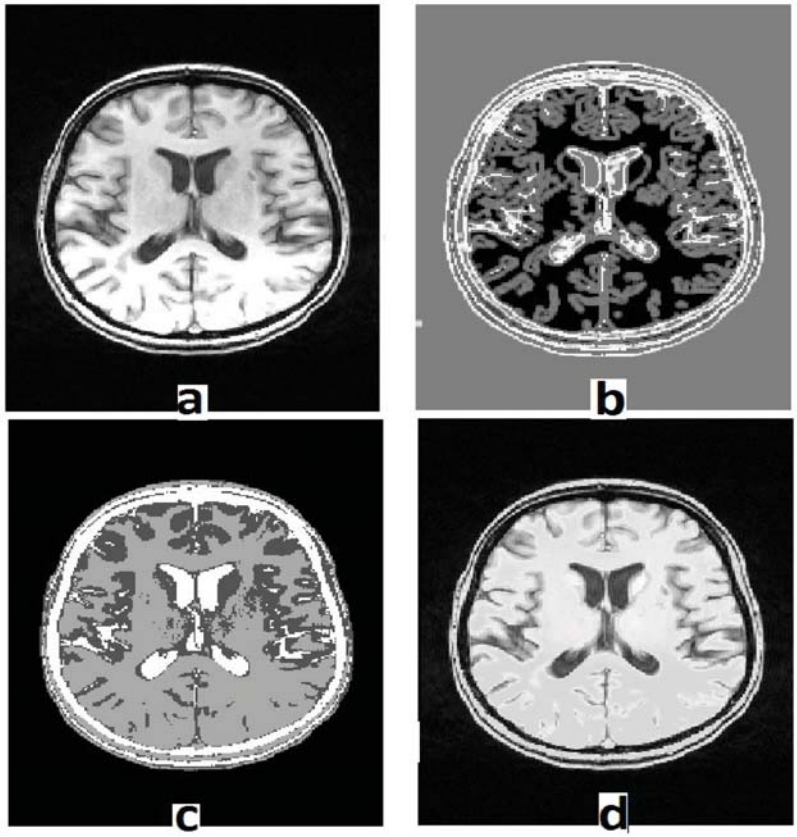


Figure 7.7: (a) Test image (b) Anatomical structural map (c) Region of interest (d) Restored image.

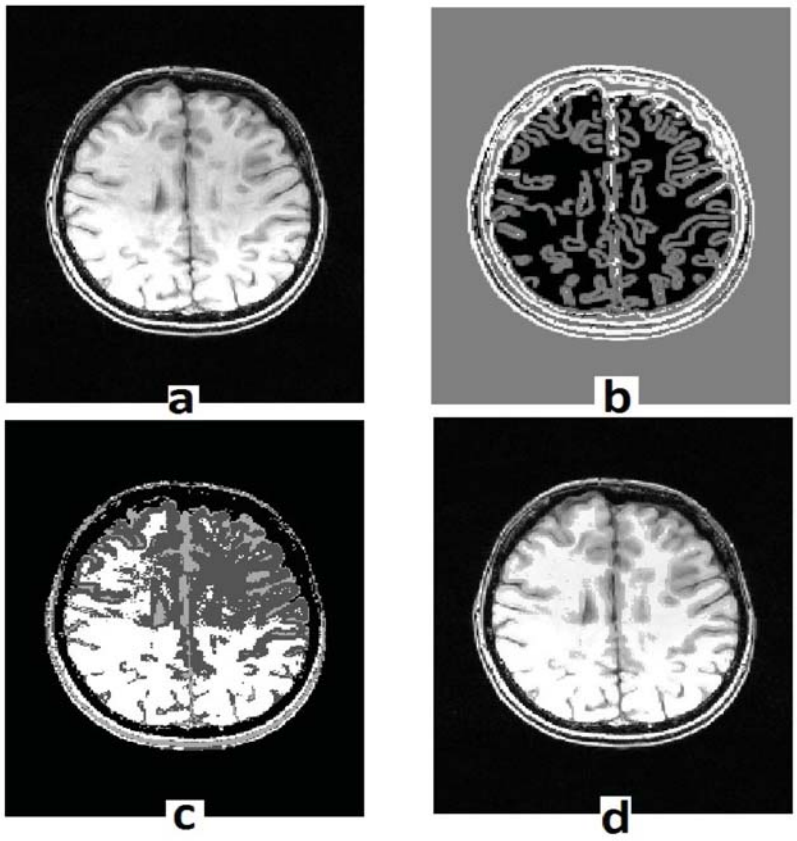


Figure 7.8: (a) Test image (b) Anatomical structural map (c) Region of interest (d) Restored image.

Chapter 8

Conclusion

This doctoral research and thesis proposes four different image processing techniques with application to brain MRI images. The first is for estimation of noise [117], the second for optimal selection of the regularization parameter [127], the third for reduction of noise [121] and the fourth is for correction of intensity inhomogeneity [128]. These proposals have several contributions.

The first layer of the MRF model is a hitherto unknown form of the Markov random field model referred to as the single layer Markov random field (SLMRF) model because there is no reference to a classical prior image model. The SLMRF extends the application of the MRF model, for the first time, to estimation of the noise level in images. The SLMRF adopts Gibbs energy, computed from a 4-neighborhood kernel, as a metric for the estimation of noise variance in brain MRI images. Comparative performance evaluation shows that the noise estimation algorithm is computationally efficient, more accurate, invariant to the presence or absence of background features in an image and is potentially immune to modeling errors inherent in current state-of-the-art algorithms. The 4-neighborhood clique kernel is a basic filter in same class as the Gaussian and wavelet filter from which state-of-the-art algorithms are derived.

There is a general framework that describes and reveals the interconnections between different algorithms that are regarded as variations of the MRF model. There is proof that confirms a long held notion that Bayesian and variational techniques are equivalent. The SLMRF was applied to develop a noise reduction algorithm that is optimal in performance on a specific type of natural images such as brain MRI images. A new approach for correction of bias fields in brain MRI images is proposed. It is based on a problem formulation which is a true reflection of presence of bias field in MRI imaging process. The proposal demonstrated excellent performance on different levels of severity of brain MRI images that were degraded by a bias field. Furthermore, this thesis pioneers the automatic generation of an anatomical structural map directly from the degraded MRI brain image.

The proposed noise estimation algorithm will improve both the process for quality assessment of images acquired from MRI machines and the performance of segmentation and denoising algorithms which requires accurate estimate of noise levels in MRI images. The proposed optimal selection of a regularization parameter in TV denoising and the SLMRF denoising method will find useful applications in MRI-based clinical trial management centers. The daily routine at the clinical trial management centers demands a high level of automation and efficiency in the processing of several thousands of MRI slice images of patients undergoing clinical trial drug treatment.

This thesis offers insight into two potential research directions. The first is the development of an integrated denoisng system which is optimal for the different classes of natural images. The second is how to develop an anatomical structural map of a single patient into subject-specific digital brain atlas. Hitherto current template-based bias field correction algorithms are guided to automatic operation by a digital brain atlas derived externally from multiple subject's MRI images. Integration of an externally derived digital brain atlas into bias field correction scheme requires the additional task of image registration.

Bibliography

[1] N. Wiest-Daessle, S. Prima, P. Coupe, S. Morrissey, and C. Barillot. Rician noise removal by non-local means filtering for low signal-to-noise ratio mri: Applications to dt-mri. In D. Metaxas, L. Axel, G. Fichtinger, and G. Szekely, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008*, volume 5242 of *Lecture Notes in Computer Science*, pages 171–179. Springer Berlin Heidelberg, 2008.

[2] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, Jan. 2001.

[3] M. E Osadebey. Simulation of realistic head geometry using radial vector representation of magnetic resonance image data. Master's thesis, Tampere University of Technology, Finland, 2009.

[4] A. C. Evans, M. Kamber, D. L. Collins, and D. MacDonald. An mri-based probabilistic atlas of neuroanatomy. In S. D. Shorvon, D. R. Fish, F. Andermann, G. M. Bydder, and H. Stefan, editors, *Magnetic Resonance Scanning and Epilepsy*, volume 264 of *NATO ASI Series*, pages 263–274. Springer US, 1994.

[5] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B Pike, and others. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1412):1293–1322, 2001.

[6] N. T Doan, J. Orban de Xivry, and B. Macq. Effect of inter-subject variation on the accuracy of atlas-based segmentation applied to human brain structures. In *Proc. SPIE*, volume 7623, pages 76231S–76231S–10, 2010.

[7] R. Gold, L. Kappos, D. L. Arnold, A. Bar-Or, G. Giovannoni, K. Selmaj, C. Tornatore, M. T. Sweetser, M. Yang, S. I. Sheikh, and others. Placebo-controlled phase 3 study of oral bg-12 for relapsing multiple sclerosis. *New England Journal of Medicine*, 367(24):2362–2362, 2012.

[8] B. Aubert-Broche, V. Fonov, K. Weier, S. Narayanan, D. L. Arnold, B. Banwell, and D. L. Collins. Is it possible to differentiate the impact of pediatric monophasic demyelinating disorders and multiple sclerosis after a first episode of demyelination? In Stanley Durrleman, Tom Fletcher, Guido Gerig, Marc Niethammer, and Xavier Pennec, editors, *Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data*, Lecture Notes in Computer Science, pages 38–48. Springer International Publishing, 2015.

[9] D. L. Arnold, G. T. Riess, P. M. Matthews, G. S. Francis, D. L Collins, C Wolfson, and J. P. Antel. Use of proton magnetic resonance spectroscopy for monitoring disease progression in multiple sclerosis. *Annals of Neurology*, 36(1):76–82, 1994.

[10] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Transactions on Medical Imaging*, 22(3):323–331, 2003.

[11] E. Lopez-Rubio and M. N Florentin-Nunez. Kernel regression based feature extraction for 3d mr image denoising. *Medical Image Analysis*, 15(4):498–513, 2011.

[12] R. Guillemaud and M. Brady. Estimating the bias field of mr images. *IEEE Transactions on Medical Imaging*, 16(3):238–251, 1997.

[13] U. Vovk, F. Pernus, and B. Likar. A review of methods for correction of intensity inhomogeneity in mri. *IEEE Transactions on Medical Imaging*, 26(3):405–421, 2007.

[14] W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. Adaptive segmentation of mri data. *IEEE Transactions on Medical Imaging*, 15(4):429–442, 1996.

[15] C. Li, C. Xu, A. W. Anderson, and J. C. Gore. Mri tissue classification and bias field estimation based on coherent local intensity clustering: A unified energy minimization framework. In J. L. Prince, D. L. Pham, and K. J. Myers, editors, *Information Processing in Medical Imaging*, volume 5636 of *Lecture Notes in Computer Science*, pages 288–299. Springer Berlin Heidelberg, 2009.

[16] J. C. Rajapakse and F. Kruggel. Segmentation of mr images with intensity inhomogeneities. *Image and Vision Computing*, 16(3):165–180, 1998.

[17] S. Roy, A. Carass, P. Bazin, and J. L Prince. Intensity inhomogeneity correction of magnetic resonance images using patches. In *SPIE Medical Imaging*, pages 79621F–79621F. International Society for Optics and Photonics, 2011.

[18] R. D Nowak. Wavelet based rician noise removal for magnetic resonance images. *IEEE Transactions on Image Processing*, 8(10):1408–1419, Oct. 1999.

[19] S. Basu, T. Fletcher, and R. Whitaker. Rician noise removal in diffusion tensor mri. In *Proceedings of the 9th international conference on Medical Image Computing and Computer-Assisted Intervention*, pages I17–125, 2006.

[20] Y. Wang and H. Zhou. Total variation wavelet-based medical image denoising. *International Journal of Biomedical Imaging*, 2006:1–6, 2006.

[21] A. Akselrod-Ballin, M. Galun, R. Basri, A. Brandt, M.J Gomori, M. Filippi, and P. Valsasina. An integrated segmentation and classification approach applied to multiple sclerosis analysis. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006*, volume 1, pages 1122–1129, Jun. 2006.

[22] G. K Rohde, A. S Barnett, P. J Basser, and C. Pierpaoli. Estimating intensity variance due to noise in registered images: Applications to diffusion tensor mri. *NeuroImage*, 26(3):673–684, 2005.

[23] L. Sendur, V. Maxim, B. Whitcher, and E. Bullmore. Multiple hypothesis mapping of functional mri data in orthogonal and complex wavelet domains. *IEEE Transactions on Signal Processing*, 53(9):3413–3426, Sept. 2005.

[24] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill. Noise and filtration in magnetic resonance imaging. *Medical Physics*, 12(5):586–591, 1985.

[25] M. J Tapiovaara and R. F Wagner. Snr and noise measurements for medical imaging: I. a practical approach based on statistical decision theory. *Physics in Medicine and Biology*, 38(1):71–92, Jan. 1993.

[26] V. S Milindkumar and P. R Deshmukh. Performance evaluation of rician noise reduction algorithm in magnetic resonance images. *Journal of Emerging Trends in Computing and Information Sciences*, 2:39–44, Feb. 2011.

[27] E. M. Haacke, R. W Brown, M. R Thompson, and R. Venkatesan. *Magnetic resonance imaging*. Wiley-Liss, 2014.

[28] P. Van de Moortele, E. J. Auerbach, C. Olman, E. Yacoub, K. Uğurbil, and S. Moeller. T 1 weighted brain images at 7 tesla unbiased for proton density, t 2âĄŐ contrast and rf coil receive b 1 sensitivity with simultaneous vessel visualization. *Neuroimage*, 46(2):432–446, 2009.

[29] J. B Arnold, J. Liow, K. A Schaper, J. J Stern, J. G Sled, D. W Shattuck, A. J Worth, M. S Cohen, R. M Leahy, J. C Mazziotta, et al. Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects. *NeuroImage*, 13(5):931–943, 2001.

[30] B. W Murphy, P. L Carson, J. H Ellis, Y. T Zhang, R. J Hyde, and T. L Chenevert. Signal-to-noise measures for magnetic resonance imagers. *Magnetic Resonance Imaging*, 11(3):425–428, 1993.

[31] M. E Brummer, R. M Mersereau, R. L Eisner, and R. R. J Lewine. Automatic detection of brain contours in mri data sets. *IEEE Transactions on Medical Imaging*, 12(2):153–166, Jun. 1993.

[32] L. Chang, G. K Rohde, and C. Pierpaoli. An automatic method for estimating noise-induced signal variance in magnitude-reconstructed magnetic resonance images. In *Medical Imaging*, pages 1136–1142. International Society for Optics and Photonics, 2005.

[33] R. M Henkelman. Measurement of signal intensities in the presence of noise in mr images. *Medical physics*, 12(2):232–233, Mar.-Apr. 1986.

[34] J. Sijbers, D. Poot, A.J den Dekker, and W. Pintjens. Automatic estimation of the noise variance from the histogram of a magnetic resonance image. *Physics in Medicine and Biology*, 52(5):1335–1348, Feb. 2007.

[35] S. Aja-Fernandez, C. Alberola-Lopez, and C. F Westin. Noise and signal estimation in magnitude mri and rician distributed images: A lmmse approach. *IEEE Transactions on Image Processing*, 17(8):1383–1398, Aug. 2008.

[36] J. Rajan, D. Poot, J. Juntu, and J. Sijbers. Noise measurement from magnitude mri using local estimates of variance and skewness. *Physics in Medicine and Biology*, 55(16):441–449, Nov. 2010.

[37] P. Coupe, J. V Manjon, E. Gedamu, D. Arnold, M. Robles, and D. L Collins. Robust rician noise estimation for mr images. *Physics in Medicine and Biology*, 14(4):483–493, Aug. 2010.

[38] J. T Heverhagen. Noise measurement and estimation in mr imaging experiments. *Radiology*, 245(3):638–639, December 2007.

[39] D. Donoho and I. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432):1200–1224, Dec. 1995.

[40] A. De Stefano, P. R. White, and W. B Collis. Training methods for image noise level estimation on wavelet components. *EURASIP Journal on Advances in Signal Processing*, 16:2400–2407, Dec. 2004.

[41] C. G Koay and P. J Basser. Analytically exact correction scheme for signal extraction from noisy magnitude mr signals. *Journal of Magnetic Resonance*, 179(2):317–322, Apr. 2006.

[42] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Transactions on Image Processing*, 7(3):310–318, Mar. 1998.

[43] B. Kimla and K. Siddiqi. Geometric heat equation and nonlinear diffusion of shapes and images. *Computer Vision and Image Understanding*, 64(3):305–322, Nov. 1996.

[44] A. Hamza, H. Krim, and G. Unal. Unifying probabilistic and variational estimation. *IEEE Signal Processing Magazine*, pages 37–47, Sept. 2002.

[45] M. Auclair-Fortier and D. Ziou. A global approach for solving evolutive heat transfer for image denoising and impainting. *IEEE Transactions on Image Processing*, 15(9):2558–2574, Sept. 2006.

[46] A. Buades, B. Coll, and J. M Morel. A non-local algorithm for image denoising. In *IEEE Conference On Computer VIsion and Pattern Recognition (CVPR05)*, volume 2, pages 60–65, 2005.

[47] F. Russo. A technique for image denoising based on adaptive piecewise linear filters and automatic parameter tuning. In *IEEE Conference on Instrumentation and Measurement Technology, IMTC 2005*, volume 2, pages 1119–1123, May 2005.

[48] A. Ben Hamza and H. Krim. Image denoising: A nonlinear robust statistical approach. *IEEE Transactions on Signal Processing*, 49(12):3045–3054, Dec. 2001.

[49] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927, 2000.

[50] J. Weickert. *Anisotropic diffusion in image processing*. B.G. Teubner, 1998.

[51] S. K Weeratunga and C. Kamath. Comparison of pde-based non-linear anistropic diffusion techniques for image denoising. In *Electronic Imaging 2003*, pages 201–212. International Society for Optics and Photonics, 2003.

[52] J. Wei. Lebesgue anisotropic image denoising. *International Journal of Imaging Systems and Technology*, 15(1):64–73, 2005.

[53] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990.

[54] J. Manjon, J. Carbonell-Caballero, J. Lull, G. Garcia-Marti, L. Marti-Bonmati, and M. Robles. Mri denoising using non-local means. *Medical Image Analysis*, 12:514–523, Feb. 2008.

[55] C. K. Chui. *An Introduction to Wavelets (Wavelet Analysis and Its Applications)*. Academic Press, San Diego, 1992.

[56] D. Donoho and I. Johnstone. Denoising by soft thresholding. *IEEE Transactions on Information Theory*, 41:613–627, 1995.

[57] X. P Zhang and M. D. Desai. Adaptive denoising based on sure risk. *IEEE Signal Processing Letters*, 5(10):265–267, 1998.

[58] F. Luisier, T. Blu, and M. Unser. A new sure approach to image denoising: Inter-scale orthornomal wavelet thresholding. *IEEE Transactions on Image Processing*, 16(3):593–6067, 2007.

[59] S. Chang, B. Yu, and M. Vetterli. Adaptive wavelet denoising for image denoising and compression. *IEEE Transactions on Image Processing*, 9:1536–1546, 2000.

[60] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, Nov. 1984.

[61] T. Chen. A markov random field model for medical image denoising. In *IEEE 2nd International Conference on Biomedical Engineering and Informatics BMEI'09*, pages 1–6, Oct. 2009.

[62] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[63] N. Varghees, M. S Manikandan, and R. Gini. Adaptive mri image denoising using total-variation and local noise estimation. In *IEEE International Conference On Advances in Engineering Science and Management(ICAESM-2012)*, pages 506–511, Mar. 2012.

[64] E. S. Gedraite and M. Hadad. Investigation on the effect of a gaussian blur in image filtering and segmentation. In *ELMAR, 2011 Proceedings*, pages 393–396, Sept 2011.

[65] S. M Monir and M. Y Siyal. Iterative adaptive spatial filtering for noise-suppression in functional magnetic resonance imaging time-series. *International Journal of Imaging Systems and Technology*, 21(3):260–270, 2011.

[66] H. Liao, F. Li, and M. K Ng. Selection of regularization parameter in total variation image restoration. *J. Opt. Soc. Am. A*, 26(11):2311–2320, Nov 2009.

[67] S. D. Babacan, R. Molina, and A. K. Katsaggelos. Parameter estimation in tv image restoration using variational distribution approximation. *IEEE Transactions on Image Processing*, 17(3):326–339, 2008.

[68] S. Ramani, Zhihao Liu, J. Rosen, J. F. Nielsen, and J. A. Fessler. Regularization parameter selection for nonlinear iterative image restoration and mri reconstruction using gcv and sure-based methods. *IEEE Transactions on Image Processing*, 21(8):3659–3672, 2012.

[69] F. Malgouyres. Minimizing the total variation under a general convex constraint for image restoration. *IEEE Transactions on Image Processing*, 11(12):1450–1456, 2002.

[70] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences*, 1(3):248–272, 2008.

[71] W. Zuo and Z. Lin. A generalized accelerated proximal gradient approach for total-variation-based image restoration. *IEEE Transactions on Image Processing*, 20(10):2748–2759, 2011.

[72] S. Arlot and A. Celisse. A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4(0):40–79, 2010.

[73] N. P. Galatsanos and A. K. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1(3):322–336, 1992.

[74] F. Catté, P. Lions, J. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.*, 29(1):182–193, February 1992.

[75] C. Knaus and M. Zwicker. Dual-domain image denoising. In *20th IEEE International Conference on Image Processing (ICIP) 2013*, pages 440–444, Sept 2013.

[76] A. Rajwade, A. Rangarajan, and A. Banerjee. Image denoising using the higher order singular value decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(4):849–862, April 2013.

[77] J. Hu, Y. Pu, X. Wu, Y. Zhang, and J. Zhou. Improved dct-based nonlocal means filter for mr images denoising. *Computational and Mathematical Methods in Medicine*, 2012:1–14, 2012.

[78] P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot. An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. *IEEE Transactions on Medical Imaging*, 27(4):425–441, 2008.

[79] Y. Zhan, M. Ding, F. Xiao, and X. Zhang. An improved non-local means filter for image denoising. In *2011 International Conference on Intelligent Computation and Bio-Medical Instrumentation (ICBMI)*, pages 31–34, Dec 2011.

[80] J. Orchard, M. Ebrahimi, and A. Wong. Efficient nonlocal-means denoising using the svd. In *15th IEEE International Conference on Image Processing, 2008*, pages 1732–1735, Oct 2008.

[81] S. P. Awate and R. T. Whitaker. Feature-preserving mri denoising: A nonparametric empirical bayes approach. *IEEE Transactions on Medical Imaging*, 26(9):1242–1255, 2007.

[82] J. Besag. On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society*, 48:259–302, 1986.

[83] S. Barnard. Stochastic stereo matching over scale. *International Journal of Computer Vision*, 3:17–32, 1989.

[84] J. S Yedidia, W. T Freeman, and Y. Weiss. Generalized belief propagation. *Advances in Neural Information Processing Systems (NIPS)*, 13:689–695, Dec. 2000.

[85] Y. Boykov, O. Veksler, and R.Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, Nov. 2001.

[86] S. Prima, N. Ayache, T. Barrick, and N. Roberts. Maximum likelihood estimation of the bias field in mr brain images: Investigating different modelings of the imaging process. In *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, MICCAI '01, pages 811–819, London, UK, UK, 2001. Springer-Verlag.

[87] W. Cui, Y. Wang, T. Lei, Y. Fan, and Y. Feng. Level set segmentation of medical images based on local region statistics and maximum a posteriori probability. *Computational and Mathematical Methods in Medicine*, 2013:1–12, 2013.

[88] C. Li, R. Huang, Z. Ding, J C Gatenby, D. N Metaxas, and J. C Gore. A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri. *IEEE Trans. on Image Process.*, 20(7):2007–2016, 2011.

[89] E. A. Vokurka, N. A. Thacker, and A. Jackson. A fast model independent method for automatic correction of intensity nonuniformity in mri data. *Journal of Magnetic Resonance Imaging*, 10(4):550–562, 1999.

[90] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig. Parametric estimate of intensity inhomogeneities applied to mri. *IEEE Transactions on Medical Imaging*, 19(3):153–165, 2000.

[91] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty. A modified fuzzy c-means algorithm for bias field estimation and segmentation of mri data. *IEEE Transactions on Medical Imaging*, 21(3):193–199, 2002.

[92] L. Liao, T. Lin, and B. Li. Mri brain image segmentation and bias field correction based on fast spatially constrained kernel clustering approach. *Pattern Recognition Letters*, 29(10):1580–1588, 2008.

[93] D. A. G. Wicks, G. J. Barker, and P. S. Tofts. Correction of intensity nonuniformity in mr images of any orientation. *Magnetic Resonance Imaging*, 11(2):183–196, 1993.

[94] L Axel, J Costantini, and J Listerud. Intensity correction in surface-coil mr imaging. *American Journal of Roentgenology*, 148(2):418–420, 1987.

[95] W. W. Brey. Correction for intensity falloff in surface coil magnetic resonance imaging. *Medical Physics*, 15(2):241, 1988.

[96] B. Johnston, M. S. Atkins, B. Mackiewich, and M. Anderson. Segmentation of multiple sclerosis lesions in intensity corrected multispectral mri. *IEEE Transactions on Medical Imaging*, 15(2):154–169, 1996.

[97] C. Han, T. S. Hatsukami, and C. Yuan. A multi-scale method for automatic correction of intensity non-uniformity in mr images. *Journal of Magnetic Resonance Imaging*, 13(3):428–436, 2001.

[98] F. Lin, Y. Chen, J. W. Belliveau, and L. L. Wald. A wavelet-based approximation of surface coil sensitivity profiles for correction of image intensity inhomogeneity and parallel imaging reconstruction. *Human Brain Mapping*, 19(2):96–111, 2003.

[99] Y. Zhuge, J. K Udupa, J. Liu, P. K Saha, and T. Iwanage. Scale-based method for correcting background intensity variation in acquired images. In *Proc. SPIE 4684, Medical Imaging 2002: Image Processing, 1103*, 2002.

[100] C. R. Meyer, P. H. Bland, and J. Pipe. Retrospective correction of intensity inhomogeneities in mri. *IEEE Transactions on Medical Imaging*, 14(1):36–41, 1995.

[101] C. Hui, Y. X Zhou, and P. Narayana. Fast algorithm for calculation of inhomogeneity gradient in magnetic resonance imaging data. *Journal of Magnetic Resonance Imaging*, 32(5):1197–1208, 2010.

[102] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of mr images of the brain. *IEEE Transactions on Medical Imaging*, 18(10):885–896, 1999.

[103] J. D. Gispert, S Reig, J Pascau, J. J. Vaquero, P Garcia-Barreno, and M Desco. Method for bias field correction of brain t1-weighted magnetic resonance images minimizing segmentation error. *Human Brain Mapping*, 22(2):133–144, 2004.

[104] J. G. Sled, A. P. Zijdenbos, and A. C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Transactions on Medical Imaging*, 17(1):87–97, 1998.

[105] B. Likar, J. B. Antoine Maintz, M. A Viergever, F Pernus, et al. Retrospective shading correction based on entropy minimization. *Journal of Microscopy*, 197(3):285–295, 2000.

[106] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.

[107] Z. Hou. A review on mr image intensity inhomogeneity correction. *International Journal of Biomedical Imaging*, 2006:1–11, 2006.

[108] R. P. Velthuizen, J. J. Heine, A. B. Cantor, H Lin, L. M. Fletcher, and L. P. Clarke. Review and evaluation of mri nonuniformity corrections for brain tumor response measurements. *Medical Physics*, 25(9):1655, 1998.

[109] G. Collewet, A. Davenel, C. Toussaint, and S. Akoka. Correction of intensity nonuniformity in spin-echo t1-weighted images. *Magnetic Resonance Imaging*, 20(4):365–373, 2002.

[110] D. Koller and N. Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[111] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer Publishing Company, 2009.

[112] W. C. Karl. *Regularization in Image Restoration and Reconstruction*. Elsevier Academic Press, Burlinton, 2nd edition, 2005.

[113] T. J. Jech. *Set theory*. Academic Press, 1978.

[114] D. J. C. Mackay. *Information Theory, Inference and learning algorithms*. Cambridge University Press, Cambridge, 2003.

[115] K. R Zalik. An efficient k-means clustering algorithm. *Pattern Recognition Letters*, 29(9):1385–1391, 2008.

[116] Y. Zhuge and J. K. Udupa. Intensity standardization simplifies brain mr image segmentation. *Computer Vision and Image Understanding*, 113(10):1095–1103, 2009.

[117] M. Osadebey, N. Bouguila, D. Arnold, and {The Alzheimer's Disease Neuroimaging Initiative}. The clique potential of markov random field in a random experiment for estimation of noise levels in 2d brain mri. *International Journal of Imaging Systems and Technology*, 23(4):304–313, 2013.

[118] A. Leon-Garcia. *Probability, statistics, and random processes for electrical engineering*. Pearson/Prentice Hall, 2008.

[119] G. R. Norman and D. L. Streiner. *Biostatistics*. Mosby, 1994.

[120] G. A. Walker. *Common statistical methods for clinical research with SAS examples*. SAS Institute, 2010.

[121] M. Osadebey, N. Bouguila, D. Arnold, and for the {Alzheimer's Disease Neuroimaging Initiative}. Four-neighborhood clique kernel: A general framework for bayesian and variational techniques of noise reduction in magnetic resonance images of the brain. *International Journal of Imaging Systems and Technology*, 24(3):224–238, 2014.

[122] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, 2004.

[123] C. Lanczos. *The Variational Principles of Mechanics*. Dover Publications, New York, 1986.

[124] J. W. Thomas. *Numerical Partial Difference Equations:Partial Difference Methods*. Springer, New York, 1995.

[125] R. Acar and C. R Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10(6):1217–1229, 1994.

[126] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990.

[127] M. Osadebey, N. Bouguila, and D. Arnold. Optimal selection of regularization parameter in total variation method for reducing noise in magnetic resonance images of the brain. *Biomedical Engineering Letters*, 4(1):80–92, 2014.

[128] M. Osadebey, N. Bouguila, and D. Arnold. *Brain MRI Intensity Inhomogeneity Correction using Region of Interest, Anatomic Structural Map and Outlier Detection*. Accepted for publication as a chpater in Elsevier edited book.

[129] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Syst., Man, Cybern.*, 9(1):62–66, 1979.

[130] C. Chou, W. Lin, and F. Chang. A binarization method with learning-built rules for document images produced by cameras. *Pattern Recognition*, 43(4):1518–1530, 2010.

[131] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, 1986.

[132] P. Bao, L. Zhang, and X. Wu. Canny edge detection enhancement by scale multiplication. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(9):1485–1490, 2005.

[133] L. G. Nyul and J. K. Udupa. On standardizing the mr image intensity scale. *Magn. Reson. Med.*, 42(6):1072–1081, 1999.

[134] F. Jaeger. *Normalization of magnetic resonance images and its application to the diagnosis of the scoliotic spine*. Logos-Verl., 2011.